



CONTRIBUTOS COMPUTACIONAIS E METODOLÓGICOS NA ESTIMAÇÃO DO ÍNDICE DE VALORES EXTREMOS

HELENA ALEXANDRA COUCEIRO FEIO DE ALMEIDA PENALVA

ORIENTADORES: Professora Doutora Maria Manuela Costa Neves Figueiredo
Professora Doutora Maria Ivette Leal de Carvalho Gomes
Professora Doutora Sandra Cristina Dias Nunes

TESE ELABORADA PARA OBTENÇÃO DO GRAU DE DOUTOR EM
MATEMÁTICA E ESTATÍSTICA

2017



CONTRIBUTOS COMPUTACIONAIS E METODOLÓGICOS NA ESTIMAÇÃO DO ÍNDICE DE VALORES EXTREMOS

HELENA ALEXANDRA COUCEIRO FEIO DE ALMEIDA PENALVA

ORIENTADORES: Professora Doutora Maria Manuela Costa Neves Figueiredo
Professora Doutora Maria Ivette Leal de Carvalho Gomes
Professora Doutora Sandra Cristina Dias Nunes

TESE ELABORADA PARA OBTENÇÃO DO GRAU DE DOUTOR EM
MATEMÁTICA E ESTATÍSTICA

JÚRI:

Presidente: Doutora Maria Margarida Branco de Brito Tavares Tomé
Professora Catedrática
Instituto Superior de Agronomia
Universidade de Lisboa.

Vogais: Doutora Maria Manuela Costa Neves Figueiredo
Professora Catedrática
Instituto Superior de Agronomia
Universidade de Lisboa;

Doutor Jorge Filipe Campinos Landerset Cadima
Professor Associado
Instituto Superior de Agronomia
Universidade de Lisboa;

Doutor Frederico Almeida Gião Gonçalves Caeiro
Professor Auxiliar
Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa;

Doutora Dora Susana Raposo Prata Gomes
Professora Auxiliar
Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa;

Doutora Maria Cristina Souto de Miranda
Professora Adjunta
Instituto Superior de Contabilidade e Administração
Universidade de Aveiro.

2017

Ao Afonso e à Maria

Agradecimentos

À minha professora Manuela, que foi uma constante fonte de inspiração e que ao longo destes anos me recebeu sempre com um sorriso luminoso e com um inabalável otimismo. Pela sua total disponibilidade e dedicação, por todas as suas sugestões e ensinamentos, pelo seu apoio, incentivo e amizade, só posso dizer um muito sentido “obrigada”. As palavras serão sempre curtas para expressar o que sinto, e o quanto aprendi e cresci com o contacto com a Professora.

À professora Ivette que me guiou nos meus primeiros passos na estatística, e que nesta fase tem sido de uma disponibilidade imensa e absolutamente imprescindível em todo o processo de reflexão e conceptualização que conduziram a este trabalho. É um enorme privilégio poder desfrutar do seu saber, do seu conhecimento e da sua experiência.

À professora Sandra que na fase inicial do trabalho foi fundamental na motivação e energia que me deu, impulsionando-me a recomeçar e a continuar, e fundamentalmente pela sua amizade e carinho ao longo destes anos.

Aos meus colegas da área científica de métodos quantitativos da ESCE pelas palavras de encorajamento, pelo carinho e amizade que me têm e pela ajuda que recebi, apoiando-me muitas vezes nas minhas responsabilidades docentes.

Aos meus filhos que me ensinam todos os dias como valem a pena todos os esforços e sacrifícios que faço.

Ao meu marido pelo amor, companheirismo e apoio incondicional. Pela confiança que sempre demonstrou ter em mim. Obrigada por estares sempre ao meu lado. Obrigada por seres quem és.

Aos meus pais pelo amor que me dedicam, pela educação que me deram, a quem eu tudo devo. Por todo o apoio e disponibilidade e por terem em muitas ocasiões me substituído no meu papel de mãe.

A minha família, em particular aos meus sogros que sempre mostraram uma disponibilidade total para me ajudarem a cuidar dos meus filhos.

Resumo

O tema central desta tese é a estimação, essencialmente em contexto semi-paramétrico, do índice de valores extremos (EVI, do inglês "*extreme value index*"), o parâmetro primordial em Teoria de Valores Extremos. Trabalha-se no âmbito de modelos de caudas pesadas, para os quais o EVI é positivo. É estudado um estimador baseado na média de Lehmer de ordem p , de números positivos, que generaliza a média aritmética ($p = 1$), e a média harmónica ($p = 0$). Dada uma amostra aleatória (X_1, \dots, X_n) e a correspondente amostra das estatísticas ordinais ascendentes $(X_{1:n}, \dots, X_{n:n})$, o clássico estimador-EVI de Hill, pode ser considerado como a média de Lehmer de ordem 1 dos k excessos, das log-observações, $V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}$, $1 \leq i \leq k < n$. Este trabalho tem como objetivo estudar o comportamento da classe de estimadores-EVI de Lehmer de ordem p . Também é estudada outra classe, a classe de Lehmer de ordem p , de viés reduzido de segunda ordem, para a qual a componente dominante do viés é eliminada. O estudo é realizado, não só, em contexto assintótico, nos respetivos níveis ótimos k , como também, em amostras de dimensão finita, com o desenvolvimento de um extenso trabalho de simulação de Monte-Carlo. É também realizada uma análise de comparação com os clássicos estimadores-EVI e com outras classes mais recentes. A um conjunto de dados reais são aplicados os vários estimadores abordados neste trabalho. De referir ainda que na obtenção dos resultados foram desenvolvidas funções no *software* R, estando em preparação um *package* R.

Palavras-chave: Caudas pesadas, Estimação semi-paramétrica, Índice de valores extremos, Redução de viés, Simulação Monte-Carlo.

Abstract

The central theme of this thesis is the estimation of the extreme value index (EVI), mainly within a semi-parametric approach. The EVI is the basic parameter in Extreme Value Theory. This work has been performed within the framework of heavy tailed parents, for which the EVI is greater than zero. A study is performed for an estimator based on Lehmer's mean of order p of a set of positive numbers, which generalizes both the arithmetic mean ($p = 1$) and the harmonic mean ($p = 0$). Given a random sample (X_1, \dots, X_n) , and the correspondent sample of ascending order statistics $(X_{1:n}, \dots, X_{n:n})$, the classic Hill EVI-estimator can be considered as the Lehmer mean of order 1 of the k log-excesses, $V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}$, $1 \leq i \leq k < n$. The main goal of this work concerns the behavior's study of Lehmer's EVI-estimator class of order p . It is also studied another class, the class of second-order reduced-bias Lehmer's EVI-estimators of order p , for which the dominant component of bias was eliminated. The study is performed, not only under an asymptotic framework, at an optimal level, k , but also for finite samples, through an extensive Monte-Carlo simulation method. A comparative analysis is also made with the classic EVI-estimators and with other more recent classes. The EVI-estimators introduced throughout this work are applied to sets of real data. It is also important to mention that, in order to achieve the desired results, it was needed to develop several functions at R *software*, being currently in development a specific R *package*.

Palavras-chave: Heavy tails, Semi-parametric estimation, Extreme value index, Reduced-bias procedures, Monte-Carlo simulation.

Conteúdo

Agradecimentos	vii
Resumo	ix
Abstract	xi
Lista de Tabelas	xv
Lista de Figuras	xvii
Lista de Símbolos e Abreviaturas	xxiii
1 Introdução	1
2 Breve Revisão Teórica em EVT	7
2.1 Alguns resultados distribucionais exatos das estatísticas ordinais	7
2.1.1 Distribuição de estatísticas ordinais	8
2.1.2 Algumas propriedades das estatísticas ordinais em modelos uni- forme, exponencial e Pareto	9
2.2 Principais resultados assintóticos em EVT	10
2.2.1 Estatísticas ordinais extremas	11
2.2.2 Caracterização dos max-domínios de atração	17
2.2.3 Variação Regular e caracterização dos domínios de atração . .	21
2.2.4 Condições de primeira e segunda ordem	25
2.2.5 Estatísticas ordinais centrais e intermédias	28
2.3 Inferência estatística em EVT	29
2.3.1 Abordagem paramétrica	31
2.3.2 Enquadramento estatístico semi-paramétrico e estimação	34
3 As Classes de Estimadores de Lehmer e outros Estimadores-EVI	39
3.1 Outros estimadores-EVI em consideração	42

3.2	Comportamento assintótico dos estimadores do EVI	46
3.3	Comparação assintótica no nível ótimo	54
3.3.1	Comparação assintótica dos estimadores-EVI L_p e H_p nos níveis ótimos	56
3.3.2	Comparação global dos estimadores-EVI nos níveis ótimos . . .	58
4	Propriedades das Classes de Estimadores de Lehmer em Amostras de Dimensão Finita	61
4.1	Padrões do valor médio e do erro quadrático médio dos estimadores-EVI	62
5	Casos de Estudo - Aplicação a Dados Reais	91
5.1	Caso de estudo: descargas médias diárias no rio Paiva	93
5.1.1	Breve análise preliminar dos dados	93
5.1.2	Teste à condição de valores extremos	94
5.1.3	A metodologia BM	95
5.1.4	A metodologia POT	97
5.1.5	A abordagem semi-paramétrica	98
5.2	Caso de estudo: área ardida em Portugal	102
5.2.1	Breve análise preliminar dos dados	102
5.2.2	Teste à condição de valores extremos	103
5.2.3	A escolha da cauda	104
5.2.4	As estimativas semi-paramétricas em estudo	104
6	Conclusões	107
6.1	Resultados	107
6.2	Trabalho Futuro	110
	Bibliografia	113
A	Trajetórias de E e RMSE	125

Lista de Tabelas

4.1	Valores de p próximos do p ótimo.	81
4.2	Estimadores com melhor desempenho no conjunto dos modelos estudados, relativamente a E	82
4.3	Estimadores com melhor desempenho no conjunto dos modelos estudados, relativamente a RMSE.	83
4.4	Valores médios simulados nos níveis ótimos e os correspondentes intervalos de confiança a 95%, relativos à família $\text{Burr}_{\xi,-0.5}$	84
4.5	Valores médios simulados nos níveis ótimos e os correspondentes intervalos de confiança a 95%, relativos à família $\text{Fréchet}_{\xi}(\rho = -1)$	85
4.6	Valores médios simulados nos níveis ótimos e os correspondentes intervalos de confiança a 95%, relativos à família $\text{Burr}_{\xi,-2}$	86
4.7	RMSE simulado de H (primeira linha) no nível ótimo e indicadores REFF juntamente com os correspondentes intervalos de confiança a 95%, relativos à família $\text{Burr}_{\xi,-0.5}$	87
4.8	RMSE simulado de H (primeira linha) no nível ótimo e indicadores REFF juntamente com os correspondentes intervalos de confiança a 95%, relativos à família $\text{Fréchet}_{\xi}(\rho = -1)$	88
4.9	RMSE simulado de H (primeira linha) no nível ótimo e indicadores REFF juntamente com os correspondentes intervalos de confiança a 95%, relativos à família $\text{Burr}_{\xi,-2}$	89
5.1	Estatísticas descritivas básicas.	94
5.2	Estimativas dos parâmetros (erros padrões em parênteses) e intervalos de 95% confiança de Log-verossimilhança de perfil (Log-L)	97
5.3	Estimativas dos parâmetros (erros padrões em parênteses) e intervalos de 95% confiança de Log-verossimilhança de perfil (Log-L)	98

5.4	Algumas estatísticas descritivas dos dados.	102
-----	---	-----

Lista de Figuras

3.1	Gráfico de $\sigma_{L_p}(\xi)/\xi$, como função de $p > 0.5$ (a) e comportamento do viés assintótico $b_{L_p}(\rho)$, para $\rho = -0.1, -0.25, -0.5$ e -1 , como função de $p \geq 0$ (b).	53
3.2	Gráfico de contornos.	57
3.3	$AREFF_{L_{p L_1}}$, como função de p , para $\rho = -3(0.1)0$ (a) e o valor de $p_{M L} = p_{M L}(\rho)$, como função ρ (b)	58
3.4	$AREFF_{L_{p_{M L} L_1}}$ e $AREFF_{H_{p_{M H} H_0}}$, como função de $\rho = -3(0.1)0$	59
3.5	Indicação dos estimadores-EVI mais eficientes calculados nos níveis ótimos.	60
4.1	Trajetórias das estimativas simuladas de E e de RMSE referentes a L para o modelo Fréchet(0.5).	64
4.2	Trajetórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Fréchet(0.5).	64
4.3	Comparação das estimativas simuladas de E e de RMSE para o modelo Fréchet(0.1).	65
4.4	Comparação das estimativas simuladas de E e de RMSE para o modelo Fréchet(0.5).	65
4.5	Comparação das estimativas simuladas de E e de RMSE para o modelo Fréchet(1).	65
4.6	Comparação das estimativas simuladas de E e de RMSE para o modelo Fréchet(2).	66
4.7	Comparação das estimativas simuladas de E e de RMSE para o modelo Fréchet(3).	66
4.8	Trajetórias das estimativas simuladas de E e de RMSE referentes a L para o modelo EV(2).	67

4.9	Trajетórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo EV(2).	68
4.10	Comparação das estimativas simuladas de E e de RMSE para o modelo EV(1).	68
4.11	Comparação das estimativas simuladas de E e de RMSE para o modelo EV(2).	69
4.12	Comparação das estimativas simuladas de E e de RMSE para o modelo EV(3).	69
4.13	Trajетórias das estimativas simuladas de E e de RMSE referentes a L para o modelo EV(0.1).	70
4.14	Trajетórias das estimativas simuladas de E e de RMSE referentes a L para o modelo EV(0.5).	70
4.15	Trajетórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo EV(0.1).	71
4.16	Trajетórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo EV(0.5).	71
4.17	Comparação das estimativas simuladas de E e de RMSE para o modelo EV(0.1).	71
4.18	Comparação das estimativas simuladas de E e de RMSE para o modelo EV(0.5).	72
4.19	Trajетórias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(0.5, -0.1).	73
4.20	Trajетórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(0.5, -0.1).	73
4.21	Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(0.5, -0.1).	74
4.22	Trajетórias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(0.1, -0.5).	74
4.23	Trajетórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(0.1, -0.5).	75
4.24	Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(0.1, -0.5).	75

4.25	Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(1, -0.5).	76
4.26	Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(3, -0.5).	76
4.27	Trajetórias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(1, -2).	77
4.28	Trajetórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(1, -2).	77
4.29	Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(1, -2).	78
4.30	Trajetórias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(0.5, -3).	78
4.31	Trajetórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(0.5, -3).	79
4.32	Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(0.1, -3).	79
4.33	Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(0.5, -3).	80
4.34	Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(1, -3).	80
5.1	Histograma (esquerda); cronograma (centro) e caixa-com-bigodes (direita).	93
5.2	Trajetória amostral da estatística E_n .	96
5.3	Histograma (direita em baixo); caixa-com-bigodes (esquerda em baixo); cronograma (esquerda em cima) e ACF (direita em cima)	96
5.4	Gráfico MRL (esquerda) e gráficos TC (centro e direita)	98
5.5	Cronograma (esquerda) com $u=180$, função de autocorrelação parcial (centro) e caixa-com-bigodes (direita).	98
5.6	Trajetórias amostrais das estatísticas G^* , R^* e W^* .	101
5.7	Trajetórias amostrais das estimativas de ξ .	101
5.8	Figuras descritivas	103
5.9	Trajetória amostral da estatística E_n .	103

5.10	Trajeto�rias amostrais das estat�sticas G^* , R^* e W^* .	104
5.11	Trajeto�rias amostrais das estimativas de ξ .	105
A.1	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a L para o modelo Fr�chet(0.1).	125
A.2	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a L para o modelo Fr�chet(1).	126
A.3	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a L para o modelo Fr�chet(2).	126
A.4	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a L para o modelo Fr�chet(3).	127
A.5	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Fr�chet(0.1).	127
A.6	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Fr�chet(1).	127
A.7	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Fr�chet(2).	128
A.8	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Fr�chet(3).	128
A.9	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a L para o modelo EV(1).	129
A.10	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a L para o modelo EV(3).	129
A.11	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a LR para o modelo EV(1).	130
A.12	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a LR para o modelo EV(3).	130
A.13	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(1, −0.1).	131
A.14	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(2, −0.1).	131
A.15	Trajeto�rias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(3, −0.1).	132

A.16 Trajetórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(1, -0.1).	132
A.17 Trajetórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(2, -0.1).	133
A.18 Trajetórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(3, -0.1).	133
A.19 Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(1, -0.1).	134
A.20 Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(2, -0.1).	134
A.21 Comparação das estimativas simuladas de E e de RMSE para o modelo Burr(3, -0.1).	134
A.22 Trajetórias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(1, -0.5).	135
A.23 Trajetórias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(3, -0.5).	135
A.24 Trajetórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(1, -0.5).	136
A.25 Trajetórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(3, -0.5).	136
A.26 Trajetórias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(0.1, -3).	137
A.27 Trajetórias das estimativas simuladas de E e de RMSE referentes a L para o modelo Burr(1, -3).	137
A.28 Trajetórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(0.1, -3).	138
A.29 Trajetórias das estimativas simuladas de E e de RMSE referentes a LR para o modelo Burr(1, -3).	138

Lista de Símbolos e Abreviaturas

v.a. variável aleatória ou variáveis aleatórias

a.a. amostra aleatória

i.i.d. independentes e identicamente distribuídas

sse se e só se

e.o. estatística ordinal ou estatísticas ordinais

$X_{i:n}$ i -ésima e.o. ascendente de X_1, X_2, \dots, X_n

$f.d.$ função de distribuição ou funções de distribuição

$f.d.p.$ função densidade de probabilidade

\frown com distribuição

f^{\leftarrow} inversa generalizada de f , ou seja, $f^{\leftarrow}(t) = \inf\{x : f(x) \geq t\}$

$U(t)$ função quantil recíproca associada à f.d. F , definida por

$$U(t) = \left(\frac{1}{1 - F} \right)^{\leftarrow}(t) = F^{\leftarrow} \left(1 - \frac{1}{t} \right), t > 1$$

$F_{i:n}$ função de distribuição da i -ésima e.o. ascendente, $X_{i:n}$

$f_{i:n}$ função densidade de probabilidade da i -ésima e.o. ascendente, $X_{i:n}$

I_A função indicatriz em A , isto é,

$$I_A(x) = \begin{cases} 0 & x \notin A \\ 1 & x \in A \end{cases}$$

$$o \quad f = o(g) \Leftrightarrow \lim_{t \rightarrow \infty} \frac{f(t)}{g(t)} = 0$$

$$O \quad f = O(g) \Leftrightarrow \lim_{t \rightarrow \infty} \frac{f(t)}{g(t)} = \lambda, \lambda \in \mathbb{R} \setminus \{0\}$$

$$\sim \quad f \sim g \Leftrightarrow \lim_{t \rightarrow \infty} \frac{f(t)}{g(t)} = 1$$

$$\xrightarrow{\mathbb{P}} \quad \text{convergência em probabilidade}$$

$$\xrightarrow{d} \quad \text{convergência em distribuição}$$

$$\stackrel{d}{\sim} \quad \text{com a mesma distribuição assintótica}$$

$$\stackrel{d}{=} \quad \text{igualdade em distribuição}$$

$$o_{\mathbb{P}} \quad X_n = o_{\mathbb{P}}(Y_n) \Leftrightarrow \frac{X_n}{Y_n} \xrightarrow{\mathbb{P}} 0, n \rightarrow \infty$$

$$O_{\mathbb{P}} \quad X_n = O_{\mathbb{P}}(Y_n) \Leftrightarrow \forall \delta > 0 \exists M_{\delta} > 0 : \mathbb{P} \left(\frac{|X_n|}{a_n} \leq M_{\delta} \right) \geq 1 - \delta, n \in \mathbb{N}$$

$$\stackrel{\mathbb{P}}{\sim} \quad \text{com o mesmo comportamento em probabilidade}$$

Capítulo 1

Introdução

A teoria de valores extremos (EVT, do inglês "*extreme value theory*") estuda o comportamento de acontecimentos que são extremos, logo pouco frequentes e que poderão vir a ser mais extremos do que aqueles que já foram observados. Neste âmbito, têm-se desenvolvido modelos que permitem realizar inferência estatística sobre esses extremos. A estatística de extremos tem sido largamente aplicada nas mais diversas áreas como a do ambiente, a hidrologia, a climatologia, a geologia e sismologia, a fiabilidade, a financeira e a da avaliação do risco, das telecomunicações, entre outras.

O primeiro livro que relata as bases teóricas em EVT, mostrando a sua importância na modelação do comportamento extremal de processos físicos e ambientais, foi *Statistics of Extremes*, Gumbel (1958). Este livro permaneceu por vários anos, e continua a ser, uma referência fundamental para as aplicações. Hoje em dia, vários livros excelentes podem ser citados com um número alargado de aplicações, por exemplo, Tiago de Oliveira (ed. 1984), e mais recentemente Embrechts *et al.* (1997), Coles (2001), Beirlant *et al.* (2004), Castillo *et al.* (2005), de Haan and Ferreira (2006), Reiss and Thomas (2007), entre outros. Recentemente surgiu um livro em português, Gomes *et al.* (2013a), com uma vasta gama de exemplos e sua resolução no *software* R.

A inferência estatística de valores extremos, a partir de observações que de alguma forma são consideradas extremas, tem como propósito modelar adequadamente a cauda da distribuição subjacente aos dados da amostra. É da maior importância a estimação de alguns parâmetros, entre os quais o índice de valores extremos, o EVI,

aqui designado também por ξ , que é a base de muitos parâmetros importantes de eventos extremos. Este parâmetro está diretamente relacionado com o peso da cauda direita da função distribuição (f.d.) F , ou seja, quanto maior o valor de ξ mais pesada é a cauda. Neste trabalho lida-se essencialmente com distribuições do tipo Pareto, i.e., com modelos de cauda pesada, para os quais o EVI é positivo. A estimação precisa de ξ é muito importante e de enorme influência na estimação de outros parâmetros de acontecimentos raros, tais como, os quantis elevados de probabilidade $1 - q$, com q pequeno, i.e., *níveis elevados* usualmente designados por *níveis de retorno*, associados aos *períodos de retorno*, $1/q$, i.e, tempo médio esperado entre excedências independentes de um nível elevado específico, a probabilidade de excedência de um determinado nível elevado ou ainda o limite superior (ou inferior) de suporte $x^F := \sup\{x : F(x) < 1\}$ (ou $x_F := \inf\{x : F(x) > 0\}$).

A abordagem paramétrica de estimação, baseada em modelos associados a resultados assintóticos no campo da EVT, foi a primeira a surgir. A teoria clássica de valores extremos diz respeito ao comportamento limite das estatísticas ordinais (e.o) extremas em amostras de dimensão n , quando $n \rightarrow \infty$, particularmente as sucessões do máximo e do mínimo. Usualmente lida-se com a cauda direita da distribuição, mas todos os resultados para máximos podem ser reformulados para mínimos. Os principais resultados assintóticos em EVT remontam a Fréchet (1927), Fisher and Tippet (1928), von Mises (1936) e Gnedenko (1943), que forneceram não só as possíveis leis limites não degeneradas de uma sucessão de valores máximos, ou mínimos, adequadamente normalizados, de uma amostra aleatória (a.a.) de dimensão n , quando $n \rightarrow \infty$, como também parcialmente caracterizaram os domínios de atração destas possíveis leis limites, i.e., determinaram quais as condições que a f.d., F , subjacente à amostra, deve satisfazer, de forma a obter-se precisamente qualquer um desses limites. A caracterização completa dos domínios de atração deve-se a de Haan (1970).

É também importante referir os resultados de Balkema and de Haan (1974) e Pickands (1975) acerca do comportamento limite dos chamados excessos acima de um determinado limiar, e a distribuição limite dos k maiores valores de uma amostra aleatória, como se pode ver em Dwass (1964).

Em finais dos anos setenta a estimação em EVT começou a ser realizada também numa abordagem semi-paramétrica, onde a estimação é realizada sob condições bastantes gerais relativamente à cauda da distribuição subjacente aos dados, e mesmo em abordagens não paramétricas. Os estimadores semi-paramétricos de EVI baseiam-se nas k e.o. superiores da amostra, ou nos excessos acima de um nível aleatório elevado, $X_{n-k:n}$. Neste contexto têm sido propostos vários estimadores de EVI. Pode-se referir os estimadores semi-paramétricos clássicos tais como, o estimador de Hill (1975) para modelos de cauda pesada (i.e., $\xi > 0$) e para $\xi \in \mathbb{R}$, o estimador de Pickands (1975), o estimador dos Momentos (Dekkers *et al.*, 1989), o estimador de Hill Generalizado (Beirlant *et al.*, 1996) e o estimador dos Momentos Mistos (Fraga Alves *et al.*, 2009).

Os estimadores semi-paramétricos supra-mencionados são consistentes para ξ , admitindo que a f.d. F pertence a um determinado domínio de atração, através de uma condição de primeira ordem, e usando uma sucessão intermédia $k \equiv k(n)$, i.e., $k \equiv k(n) \rightarrow \infty$ e $\frac{k(n)}{n} \rightarrow 0$ quando $n \rightarrow \infty$. Adicionando ainda uma condição de segunda ordem, que regula a convergência da condição de primeira ordem, é possível também garantir-se a convergência em distribuição, desses estimadores, para a normal. Seja $\hat{\xi}^\bullet(k)$ qualquer um dos estimadores referidos e $A(t)$ uma função convergente para zero quando $t \rightarrow \infty$, que mede a velocidade de convergência na condição de primeira ordem. Se $\sqrt{k}A(n/k) \rightarrow \lambda$ finito, quando $n \rightarrow \infty$, então $\sqrt{k} \left(\hat{\xi}^\bullet(k) - \xi \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N} \left(\lambda b^\bullet, \sigma^{\bullet 2} \right)$. A $b^\bullet A(n/k)$ e $\sigma^{\bullet 2}/k$ dá-se o nome respetivamente de *viés assintótico* e *variância assintótica* de $\hat{\xi}^\bullet(k)$.

A maioria dos estimadores semi-paramétricos clássicos exibem a mesma dificuldade: uma variância elevada para limiares elevados, $X_{n-k:n}$, i.e., para valores pequenos de k ; e viés elevado para limiares baixos, i.e., para valores elevados de k . Consequentemente, a escolha do limiar deverá ter em conta a minimização do erro quadrático médio (MSE, do inglês "*mean square error*") e constitui um problema difícil em EVT, um tema que se encontra ainda em desenvolvimento. Este problema tem levado os investigadores a construir novos estimadores de EVI, de viés reduzido

de segunda ordem (SORB, do inglês "*second order reduced bias*"). Esta abordagem tem sido considerada essencialmente para caudas pesadas e a ideia base é encontrar forma de eliminar a componente dominante do viés, $b \bullet A(n/k)$. Nesses trabalhos, foram desenvolvidos estimadores de EVI em que se conseguiu reduzir o viés assintótico mas com o custo de aumentar a variância. Mais recentemente o problema de "trade-off" entre viés e variância tem vindo a ser ultrapassado, para modelos de caudas pesadas, com a introdução dos denominados estimadores de viés-reduzido (RB, do inglês "*reduced-bias*") com variância-mínima (MVRB, do inglês "*minimum-variance reduced-bias*"). Para uma revisão geral das diversas abordagens ver Gomes *et al.* (2008c), Beirlant *et al.* (2012) e Gomes and Guillou (2015).

Neste trabalho realiza-se um estudo sobre uma classe de estimadores de EVI, a classe de estimadores de Lehmer de ordem p , L_p , e a correspondente classe de estimadores de Lehmer de viés reduzido de segunda ordem, L_p^{RB} , comparando-as com outros estimadores-EVI, inclusivamente com classes de estimadores mais recentes e promissoras. De realçar também que na obtenção dos resultados foram desenvolvidas várias funções no *software* R, quer para o vasto estudo de simulação realizado, quer para calcular as diferentes estimativas nos casos de aplicação a dados reais.

No Capítulo 2 apresenta-se uma breve revisão dos resultados mais relevantes em EVT, considerando que se trabalha com variáveis aleatórias (v.a) independentes e identicamente distribuídas (i.i.d.). Começa-se por caracterizar o comportamento exato de uma e.o. Dado que o comportamento assintótico de uma e.o., $X_{k:n}$, depende da forma como $k \equiv k(n)$ se relaciona com n , classificam-se as e.o. em extremas, centrais ou intermédias, e estuda-se o seu comportamento assintótico. Nas e.o. extremas dá-se especial destaque à distribuição de Valores Extremos e à caracterização dos domínios de atração. Essa caracterização é feita também através da classe de funções de variação regular generalizada. Apresentam-se as condições de primeira e segunda ordem para modelos de cauda pesada. Os modelos de cauda pesada utilizados neste trabalho satisfazem as condições até à segunda ordem e pertencem à classe de Hall-

Welsh, cujo desenvolvimento, de segunda ordem, da cauda direita é da forma

$$\overline{F}(x) = (x/C)^{-1/\xi} \left(1 + \beta(x/C)^{\rho/\xi} / \rho + o(x^{\rho/\xi}) \right), \quad \text{quando } x \rightarrow \infty,$$

$C > 0$, $\beta \neq 0$ e $\rho < 0$. Neste capítulo é também realizada uma breve revisão dos métodos de estimação em contexto paramétrico, bem como semi-paramétrico.

No Capítulo 3, introduz-se a classe de estimadores-EVI Lehmer de ordem p , L_p e é realizado o estudo do seu comportamento assintótico. Também esta classe de estimadores apresenta as mesmas dificuldades já indicadas para os estimadores semi-paramétricos. Assim introduz-se um estimador de viés reduzido, L_p^{RB} , construído com base nesta classe, e que constitui uma generalização do estimador MVRB, já referido. São apresentados vários estimadores-EVI existentes na literatura, com o objetivo de serem usados nas comparações do comportamento assintótico, a nível ótimo, com as duas classes de estimadores agora introduzidas.

No Capítulo 4 é desenvolvido um extenso trabalho de simulação de Monte-Carlo para estudar o comportamento das classes de estimadores de Lehmer em amostras de dimensão finita. Também, neste âmbito, é realizada uma comparação com os outros estimadores-EVI.

No Capítulo 5 ilustra-se, com conjuntos de dados reais, a estimação de EVI em contexto paramétrico e também através dos estimadores semi-paramétricos estudados ao longo do trabalho.

Finalmente as conclusões do presente estudo serão apresentadas no Capítulo 6.

Capítulo 2

Breve Revisão Teórica em EVT

Neste capítulo apresentam-se alguns dos principais resultados e definições em EVT para v.a. i.i.d. Na Secção 2.1 caracteriza-se o comportamento exato de e.o. e apresentam-se algumas propriedades das e.o. em modelos uniforme, exponencial e Pareto. Na Secção 2.2 caracteriza-se o comportamento assintótico das e.o. centrais, intermédias e extremas. Salienta-se a distribuição limite não degenerada, G , e as suas possíveis formas, do máximo linearmente normalizado. Nesta Secção também se apresenta, recorrendo à teoria das funções de variação regular generalizada e à função quantil recíproca, uma caracterização dos domínios de atração para máximos da distribuição limite não degenerada, G . Por último, na Secção 2.3 faz-se uma breve descrição das abordagens de estimação utilizadas na inferência estatística de valores extremos.

2.1 Alguns resultados distribucionais exatos das estatísticas ordinais

Para um estudo elementar da teoria distribucional de e.o., veja-se Arnold *et al.* (1992; 2008), David and Nagaraja (2003) e Gomes *et al.* (2013a). Nestes livros encontram-se os resultados e demonstrações que se indicam a seguir.

2.1.1 Distribuição de estatísticas ordinais

Sejam X_1, \dots, X_n n v.a. i.i.d. com f.d. F e sejam $X_{1:n} \leq \dots \leq X_{n:n}$ as respectivas e.o. ascendentes. Tem-se para $x \in \mathbb{R}$

$$F_{X_{i:n}}(x) = \mathbb{P}(\text{ pelo menos } i \text{ das } n \text{ observações } \leq x) = \sum_{k=i}^n \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k}. \quad (2.1)$$

Assim no caso do mínimo da amostra tem-se,

$$F_{X_{1:n}}(x) = 1 - (1 - F(x))^n \quad (2.2)$$

e no caso do máximo,

$$F_{X_{n:n}}(x) = F^n(x). \quad (2.3)$$

Se a v.a. X for absolutamente contínua, a função densidade de probabilidade (f.d.p.) de $X_{i:n}$ é dada por,

$$f_{X_{i:n}}(x) = \frac{1}{B(i, n-i+1)} [F(x)]^{i-1} [1 - F(x)]^{n-i} f(x), \quad (2.4)$$

onde $B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$, $a > 0$, $b > 0$, é a função Beta completa.

Também é possível obter-se a f.d. do par aleatório $(X_{i:n}, X_{j:n})$, $1 \leq i < j \leq n$,

$$F_{X_{i:n}, X_{j:n}}(x, y) = \begin{cases} \sum_{k=j}^n \sum_{l=i}^k \frac{n!}{l!(k-l)!(n-k)!} F^l(x) [F(y) - F(x)]^{k-l} [1 - F(y)]^{n-k}, & x < y \\ F_{X_{j:n}}(y), & x \geq y \end{cases} \quad (2.5)$$

e a f.d.p. conjunta do par aleatório $(X_{i:n}, X_{j:n})$, $1 \leq i < j \leq n$,

$$f_{X_{i:n}, X_{j:n}}(x, y) = \frac{1}{B(i, j-i)B(j, n-j+1)} F^{i-1}(x) [F(y) - F(x)]^{j-i-1} [1 - F(y)]^{n-j} f(x)f(y), \quad x < y. \quad (2.6)$$

2.1.2 Algumas propriedades das estatísticas ordinais em modelos uniforme, exponencial e Pareto

Considere-se o conceito de função quantil recíproca associada à f.d. F .

Definição 2.1.1 (Função quantil recíproca associada à f.d. F). *Seja F uma f.d. de uma v.a. X e denote-se por F_X^{\leftarrow} a respetiva inversa generalizada,*

$$F_X^{\leftarrow}(t) := \inf \{x : F_X(x) \geq t\}. \quad (2.7)$$

A correspondente função quantil recíproca associada à f.d. F de uma v.a. X é definida por,

$$U_X(t) := \left(\frac{1}{1-F} \right)^{\leftarrow}(t) = F_X^{\leftarrow} \left(1 - \frac{1}{t} \right), \quad t \in [1, \infty[. \quad (2.8)$$

Observação 2.1.1. *Tem-se obviamente,*

$U_X(\infty) := \lim_{t \rightarrow \infty} U_X(t) = F_X^{\leftarrow}(1) = \inf \{x : F_X(x) \geq 1\} = \sup \{x : F_X(x) < 1\} = x^F$, o limite superior do suporte.

Teorema 2.1.1 (Rényi 1953). *Sejam $E_{1:n}, \dots, E_{n:n}$ as e.o. ascendentes associadas a uma a.a. de dimensão n com distribuição exponencial unitária. Então, com $E_{0:n} \equiv 0$,*

1. *os espaçamentos $(E_{i:n} - E_{i-1:n})$ e $(E_{j:n} - E_{j-1:n})$ são independentes para $i \neq j$ com $i, j = 1, \dots, n$;*
2. *$(n-i+1)(E_{i:n} - E_{i-1:n}) \cap \text{Exp}(1)$.*

Observação 2.1.2. *O ponto 2. do Teorema anterior pode reescrever-se como,*

$$(n-i+1)(E_{i:n} - E_{i-1:n}) \stackrel{d}{=} E_i, \quad 1 \leq i \leq n, \quad (2.9)$$

que permite obter a conhecida representação de Rényi,

$$E_{i:n} = \sum_{l=1}^i (E_{l:n} - E_{l-1:n}) \stackrel{d}{=} \sum_{l=1}^i \frac{E_l}{n-l+1}, \quad 1 \leq i \leq n, \quad (2.10)$$

e a seguinte relação,

$$E_{j:n} - E_{i:n} \stackrel{d}{=} E_{j-i:n-i}, \quad 1 \leq i < j \leq n. \quad (2.11)$$

Teorema 2.1.2. *Sejam $U_{1:n}, \dots, U_{n:n}$ as e.o. ascendentes associadas a uma a.a. de dimensão n com distribuição uniforme em $(0, 1)$, e sejam $X_{1:n}, \dots, X_{n:n}$ as e.o. ascendentes associadas a uma a.a. de dimensão n com f.d. $F(\cdot)$ arbitrária. Tem-se,*

$$F^{\leftarrow}(U_{i:n}) \stackrel{d}{=} X_{i:n}. \quad (2.12)$$

Teorema 2.1.3. *Sejam X_1, \dots, X_n n v.a. i.i.d. com f.d. $F(\cdot)$, arbitrária, Y_1, \dots, Y_n v.a. i.i.d. com distribuição de Pareto unitária e E_1, \dots, E_n n v.a. i.i.d. com distribuição exponencial unitária. Então,*

1. $\ln Y_{i:n} \stackrel{d}{=} E_{i:n};$
2. $X_{i:n} \stackrel{d}{=} U_X(Y_{i:n});$
3. $\frac{Y_{j:n}}{Y_{i:n}} \stackrel{d}{=} Y_{j-i:n-i}.$

2.2 Principais resultados assintóticos em EVT

O comportamento assintótico da e.o. $X_{k:n}$ depende da relação existente entre $k \equiv k(n)$ e n . Por conseguinte, é possível classificar em três grupos as e.o.:

- Se $k(n) \xrightarrow{n \rightarrow \infty} \infty$ e $\frac{k(n)}{n} \xrightarrow{n \rightarrow \infty} \alpha \in]0, 1[$, $X_{k(n):n}$ diz-se uma e.o. central.
- Se $k(n) \xrightarrow{n \rightarrow \infty} \infty$ e $\frac{k(n)}{n} \xrightarrow{n \rightarrow \infty} 0$, $X_{k(n):n}$ diz-se uma e.o. intermédia inferior e $X_{n-k(n):n}$ uma e.o. intermédia superior.
- Se $k(n) = r$ ou $k(n) = n - r$ onde r é um valor fixo, diz-se que $X_{k(n):n}$ é uma e.o. extremal inferior ou superior, respetivamente.

No caso das e.o. intermédias superiores, $X_{n-k(n):n}$, considera-se que $k \equiv k(n)$ é uma sucessão intermédia, i.e., uma sucessão de valores inteiros entre 1 e n que verifica,

$$k \equiv k(n) \xrightarrow{n \rightarrow \infty} \infty, \quad \frac{k(n)}{n} \xrightarrow{n \rightarrow \infty} 0, \quad \text{i.e.} \quad k = o(n). \quad (2.13)$$

2.2.1 Estatísticas ordinais extremas

A teoria clássica de valores extremos diz respeito ao comportamento limite do máximo, $X_{n:n}$, ou do mínimo, $X_{1:n}$, quando $n \rightarrow \infty$, de uma sucessão de v.a. $\{X_n\}_{n \geq 1}$ i.i.d. (ou possivelmente estacionárias e fracamente dependentes) com f.d desconhecida F . Considere-se assim $X_{1:n} \leq \dots \leq X_{n:n}$ as e.o. ascendentes associadas à amostra (X_1, \dots, X_n) . Usualmente lida-se com a cauda direita da distribuição,

$$\bar{F}(x) := 1 - F(x), \quad (2.14)$$

para um x elevado, mas todos os resultados para máximos podem ser reformulados para mínimos, tendo em consideração que $\min_{1 \leq i \leq n}(X_i) = -\max_{1 \leq i \leq n}(-X_i)$, ou de uma forma mais geral, utilizando a relação que se verifica entre as e.o. superiores e as inferiores, $X_{i:n} = -Y_{n-i+1:n}$, $1 \leq i \leq n$, com $Y_i = -X_i$. É bem conhecido que nas condições referidas, a distribuição do máximo, $X_{n:n}$, é $F^n(\cdot)$. Contudo esta distribuição é de pouca relevância na prática pois, por um lado, sendo F desconhecida e podendo ser mal especificada, pode levar a grandes erros na distribuição do máximo. Por outro lado, fazendo $n \rightarrow \infty$, a f.d. de $X_{n:n}$,

$$F_{X_{n:n}}(x) = F^n(x) \xrightarrow{n \rightarrow \infty} \begin{cases} 0, & x < x^F \\ 1, & x \geq x^F \end{cases},$$

é degenerada, tendo por isso pouco interesse. Como se irá referir, a solução encontrada consistiu em obter a distribuição limite de $X_{n:n}$ (ou de $X_{1:n}$), conveniente e linearmente normalizado, quando $n \rightarrow \infty$, ou seja a distribuição limite de $\frac{X_{n:n} - b_n}{a_n}$ $\left(\frac{X_{1:n} - b_n}{a_n}\right)$, onde b_n representa uma mudança de localização e a_n de escala. Assim, o principal problema que se colocou em EVT consistiu em determinar para que

f.d. de X existem sucessões $\{a_n\}_{n \geq 1} \in \mathbb{R}^+$ e $\{b_n\}_{n \geq 1} \in \mathbb{R}$, tais que

$$\mathbb{P} \left(\frac{X_{n:n} - b_n}{a_n} \leq x \right) = F^n(a_n x + b_n) \xrightarrow{n \rightarrow \infty} G(x), \quad (2.15)$$

para todos os pontos de continuidade x de G e onde G é uma f.d. não degenerada. Este problema tem duas vertentes, a primeira é identificar todas as possíveis leis limite não degeneradas, G , de uma sucessão de valores máximos de uma amostra aleatória de dimensão n , adequadamente normalizados, quando $n \rightarrow \infty$, e a segunda é a caracterização dos domínios de atração das possíveis leis limite para máximos, i.e., quais as condições que a distribuição F deve satisfazer de forma a obter-se precisamente essa distribuição limite.

A primeira parte do problema foi resolvida por Fréchet (1927), Fisher and Tippett (1928), von Mises (1936) e rigorosamente demonstrada por Gnedenko (1943) e por de Haan (1970). Relativamente à segunda parte do problema, foram von Mises (1936) e Gnedenko (1943) que parcialmente caracterizaram os domínios de atração destas possíveis leis limite, ficando-se a dever a de Haan (1970) a sua completa caracterização. Estes autores forneceram condições para a existência de sucessões $\{a_n\}_{n \geq 1} \in \mathbb{R}^+$ e $\{b_n\}_{n \geq 1} \in \mathbb{R}$, constantes normalizadoras (chamadas constantes de atração) tais que (2.15) é válida. Se tal se verificar diz-se que F pertence ao domínio de atração para máximos da f.d G e escreve-se $F \in \mathcal{D}_{\mathcal{M}}(G)$.

Antes de se determinar que formas pode assumir a função G , é importante perceber dois conceitos.

Definição 2.2.1 (Distribuições do mesmo tipo). *Duas f.d. F_1 e F_2 dizem-se do mesmo tipo se existem constantes reais $a > 0$ e b tais que,*

$$F_2(x) = F_1(ax + b).$$

Ou seja duas distribuições do mesmo tipo são a mesma a menos de parâmetros de localização e escala.

Teorema 2.2.1 (Convergência de tipos de Khinchine). *Sejam F_1 e F_2 duas f.d. não degeneradas.*

1. Se existirem f.d. F_n e sucessões reais $a_n > 0$, $a'_n > 0$, b_n e b'_n tais que $F_{X_{n:n}}(a_n x + b_n) \xrightarrow{n \rightarrow \infty} F_1(x)$ e $F_{X_{n:n}}(a'_n x + b'_n) \xrightarrow{n \rightarrow \infty} F_2(x)$, então quando $n \rightarrow \infty$,

$$\frac{a'_n}{a_n} \rightarrow a, \quad \frac{b'_n - b_n}{a_n} \rightarrow b \text{ e } F_2(x) = F_1(ax + b).$$

Isto é, (a'_n, b'_n) é assintoticamente equivalente a (a_n, b_n) e F_1 e F_2 são distribuições do mesmo tipo.

2. Inversamente, se $\frac{a'_n}{a_n} \rightarrow a > 0$, $\frac{b'_n - b_n}{a_n} \rightarrow b$ e $F_{X_{n:n}}(a_n x + b_n) \xrightarrow{n \rightarrow \infty} F_1(x)$ então,

$$F_{X_{n:n}}(a'_n x + b'_n) \xrightarrow{n \rightarrow \infty} F_1(ax + b).$$

Este teorema veio demonstrar que as sucessões a_n e b_n não são únicas e que se uma f.d. F pertence a um determinado domínio de atração, então não pertence a mais nenhum. Veja-se então quais as distribuições limite para o máximo normalizado de uma amostra aleatória.

Teorema 2.2.2 (Teorema dos tipos extremais: Gnedenko, 1943). *Seja F uma f.d. que pertence ao domínio de atração para máximos de uma f.d G , não degenerada. Então a distribuição G pertence a um dos três tipos de distribuições de valores extremos:*

$$\text{Tipo I: Gumbel} \quad \Lambda(x) = \exp(-\exp(-x)), \quad x \in \mathbb{R}. \quad (2.16a)$$

$$\text{Tipo II: Fréchet} \quad \Phi_\alpha(x) = \exp(-x^{-\alpha}), \quad x > 0, \alpha > 0. \quad (2.16b)$$

$$\text{Tipo III: Max-Weibull} \quad \Psi_\alpha(x) = \exp(-(-x)^\alpha), \quad x < 0, \alpha > 0. \quad (2.16c)$$

Observação 2.2.1. *Este teorema afirma que os três tipos de distribuições de valores extremos são as únicas possíveis leis limite para a distribuição do máximo normalizado independentemente da f.d. F subjacente à a.a.*

Estas três famílias de distribuições são também denominadas de max-estáveis. Veja-se o significado deste conceito.

Definição 2.2.2. *Uma distribuição G diz-se max-estável se existem constantes $\alpha_k > 0$*

e β_k tais que,

$$G^k(x) = G(\alpha_k x + \beta_k), \forall_k.$$

Admitindo a validade de (2.15) e sendo $X_{nk:nk}$ o máximo de uma a.a. de dimensão nk ou o máximo de k máximos de k amostras aleatórias de dimensão n , tem-se

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{X_{nk:nk} - b_{nk}}{a_{nk}} \leq x \right] = G(x) \Leftrightarrow$$

$$\lim_{n \rightarrow \infty} F^{nk}(a_{nk}x + b_{nk}) = G(x) \Leftrightarrow$$

$$\lim_{n \rightarrow \infty} F^n(a_{nk}x + b_{nk}) = G^{\frac{1}{k}}(x).$$

Assim pelo teorema de convergência de tipos de Khinchine, a distribuição G e $G^{\frac{1}{k}}$ são do mesmo tipo, i.e., existem (α_k, β_k) , tais que $G^{\frac{1}{k}}(\alpha_k x + \beta_k) = G(x) \Leftrightarrow G^k(x) = G(\alpha_k x + \beta_k)$ e portanto a distribuição G é max-estável. Pode-se assim referir que se existe uma distribuição limite para o máximo normalizado, então essa distribuição é max-estável e que a distribuição limite do máximo de k máximos é da mesma família da distribuição limite do máximo.

O teorema seguinte formaliza o resultado anterior.

Teorema 2.2.3. *A função G é uma função distribuição de valores extremos se e só se G for max-estável.*

O Teorema de Gnedenko admite a validade de (2.15), mas no entanto existem sucessões de v.a. $\{X_n\}_{n \geq 1}$ para as quais não é possível encontrar sucessões de constantes $\{a_n\}_{n \geq 1} \in \mathbb{R}^+$ e $\{b_n\}_{n \geq 1} \in \mathbb{R}$ de forma a que tal aconteça, como se pode verificar atendendo aos dois teoremas seguintes.

Teorema 2.2.4. *Se $\{X_n\}_{n \geq 1}$ é uma sucessão de v.a. i.i.d. com f.d. comum $F(\cdot)$ tal que $F(x^F -) < 1$, $x^F < \infty$, então não é possível normalizar $X_{n:n}$ de modo a obter a convergência em distribuição para uma v.a. não degenerada.*

Teorema 2.2.5. *Seja $F(x)$ uma f.d. tal que para uma sucessão $x_1 < \dots < x_n < \dots$ de reais com limite finito ou infinito,*

$$\frac{1 - F(x_{n-})}{1 - F(x_n)} \geq 1 + \beta, \quad \beta > 0.$$

Então F não pertence ao domínio de atração de nenhuma das leis limites de máximos.

Observação 2.2.2. Por exemplo não é possível normalizar o máximo de v.a.'s com f.d. de Poisson de modo a obter-se uma distribuição limite não degenerada.

As três famílias de distribuições, indicadas em (2.16), foram unificadas por von Mises (1936, 1964) e Jenkinson (1955) através da f.d.,

$$G(x) \equiv EV_{\xi}(x) = \begin{cases} \exp \left\{ -[1 + \xi x]^{-1/\xi} \right\}, & 1 + \xi x > 0, \text{ se } \xi \neq 0 \\ \exp \{ -\exp[-x] \}, & x \in \mathbb{R}, \text{ se } \xi = 0, \end{cases} \quad (2.17)$$

onde $\xi \in \mathbb{R}$ é o parâmetro de forma da distribuição EV_{ξ} , o EVI.

A f.d. EV_{ξ} incorpora, de facto, as três famílias:

- **A família Gumbel**, corresponde a $\xi = 0$ (limite por continuidade quando $\xi \rightarrow 0^-$ ou $\xi \rightarrow 0^+$),

$$\Lambda(x) = EV_0(x) = \exp(-\exp(-x)), \quad x \in \mathbb{R}, \xi = 0;$$

- **A família Fréchet**, corresponde a $\xi > 0$,

$$\Phi_{\alpha}(x) = EV_{1/\alpha}(\alpha(x-1)) = \exp(-x^{-\alpha}), \quad x > 0, \xi = 1/\alpha > 0;$$

- **A família Weibull**, corresponde a $\xi < 0$,

$$\Psi_{\alpha}(x) = EV_{-1/\alpha}(\alpha(x+1)) = \exp(-(-x)^{\alpha}), \quad x < 0, \xi = -1/\alpha < 0.$$

Em toda a análise de valores extremos, $\xi \in \mathbb{R}$, é um parâmetro primordial, que está diretamente relacionado com o peso da cauda direita da f.d. F , definida em (2.14), ou seja, quanto maior o valor de ξ mais pesada é a cauda direita.

1. Para $\xi = 0$, a cauda direita é do tipo exponencial, o *limite superior de suporte* de F , x^F , pode ser finito ou infinito. Está-se no domínio de atração para máximos

da Gumbel, que constitui o limite associado a muitas distribuições conhecidas tais como a distribuição normal, lognormal, exponencial e gama.

2. Para $\xi > 0$, a cauda direita é pesada, do tipo polinomial negativa e F tem *limite superior de suporte* infinito. Os momentos de ordem r são infinitos se $r \geq 1/\xi$. A distribuição de Pareto e a Cauchy são exemplos de distribuições que pertencem ao domínio de atração para máximos da f.d. Fréchet.
3. Para $\xi < 0$, a cauda direita é curta, e F tem *limite superior de suporte* finito, $x^F = -1/\xi < +\infty$. Pode-se referir a distribuição uniforme como um exemplo de uma distribuição que pertence ao domínio de atração para máximos da distribuição Weibull.

Embora possa parecer que a família de distribuições limite em (2.17), apresenta algumas dificuldades de aplicação devido ao desconhecimento das constantes de normalização $\{a_n\}$ e $\{b_n\}$, este limite pode ser considerado, para um n suficientemente grande, como

$$\mathbb{P}\left(\frac{X_{n:n} - b_n}{a_n} \leq x\right) \approx EV_\xi(x).$$

Por outro lado,

$$\mathbb{P}(X_{n:n} \leq x) = F^n(x) \approx EV_\xi\left(\frac{x - b_n}{a_n}\right) = EV_\xi^*(x),$$

onde EV_ξ^* é também uma f.d. da mesma família de EV_ξ . Assim as constantes normalizadoras podem ser incorporadas numa versão da f.d. EV_ξ que considere os parâmetros de localização e escala, $\lambda \in \mathbb{R}$ e $\delta \in \mathbb{R}^+$, respetivamente, $EV_\xi(x; \lambda, \delta) \equiv EV_\xi((x - \lambda)/\delta)$. Uma vez que estes parâmetros têm de ser estimados, é irrelevante na prática que se desconheçam as constantes de normalização.

Alternativamente a considerar apenas o valor máximo de uma a.a. como um valor extremo, pode-se definir também todas as observações, X_i , acima de um específico nível elevado ou limiar, u , como sendo valores extremos. As diferenças $X_i - u$, são chamadas de excessos acima desse limiar. Balkema and de Haan (1974) e Pickands (1975) provaram que a f.d. Generalizada de Pareto, GP , é a distribuição limite dos excessos acima de um limiar elevado, adequadamente escalados.

Teorema 2.2.6. *Seja X uma v.a. com f.d. $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$. Então para u suficientemente elevado, a f.d. de $Y = (X - u) | X > u$ é aproximadamente a f.d. Generalizada Pareto (GP),*

$$GP_{\xi}(y) = \begin{cases} 1 - (1 + \xi y/\tilde{\delta})^{-1/\xi}, & y > 0, (1 + \xi y/\tilde{\delta}) > 0, \text{ se } \xi \neq 0, \\ 1 - \exp(-y/\tilde{\delta}), & y > 0, \text{ se } \xi = 0, \end{cases} \quad (2.18)$$

onde ξ é o parâmetro de forma, igual ao da correspondente distribuição EV, e $\tilde{\delta} = \delta + \xi(u - \lambda)$ o parâmetro de escala onde λ é o parâmetro de localização da f.d. EV. O recíproco deste resultado é também verdadeiro.

Também se pode considerar o comportamento das k e.o. de topo na caracterização do comportamento extremal de uma variável.

Teorema 2.2.7 (Distribuição assintótica das k e.o. superiores). *A f.d. F pertence ao domínio de atração da f.d. EV se e só se, para k fixo, o vetor aleatório k -dimensional, adequadamente normalizado pelas constantes $\{a_n\} \in \mathbb{R}^+$ e $\{b_n\} \in \mathbb{R}$,*

$$\left(\frac{X_{n:n} - b_n}{a_n}, \dots, \frac{X_{n-k+1:n} - b_n}{a_n} \right),$$

tem distribuição limite, quando $n \rightarrow \infty$, com f.d.p conjunta,

$$g(w_1, \dots, w_k) = EV_{\xi}(w_k) \prod_{i=1}^k \frac{ev_{\xi}(w_i)}{EV_{\xi}(w_i)}, \quad w_1 > \dots > w_k, \quad (2.19)$$

com $EV_{\xi}(w)$ definida em (2.17) e onde $ev_{\xi}(w) = \frac{\partial EV_{\xi}(w)}{\partial w}$ é a f.d.p. do modelo EV. Este modelo é usualmente denominado por modelo multivariado-EV $_{\xi}$, também conhecido como processo extremal, e foi estudado em Dwass (1964).

Observação 2.2.3. *Note-se que o vetor $(X_{n:n}, \dots, X_{n-k+1:n})$ pode ser normalizado da mesma forma que o máximo.*

2.2.2 Caracterização dos max-domínios de atração

Em EVT é fundamental a caracterização dos domínios de atração, i.e., dada uma distribuição limite, pretende-se identificar o conjunto de distribuições F para as quais a

f.d. do máximo convenientemente normalizado converge para essa distribuição limite, ou equivalentemente, determinar quais as condições que a f.d. F deve satisfazer de forma a existirem sucessões $\{a_n\}_{n \geq 1} \in \mathbb{R}^+$ e $\{b_n\}_{n \geq 1} \in \mathbb{R}$ para as quais se verifique a convergência em distribuição em (2.15).

Foi von Mises (1936) quem estabeleceu uma condição suficiente, que garante que uma dada distribuição F pertence a um determinado domínio de atração. Reveja-se em primeiro lugar a Definição 2.1.1 de função quantil recíproca associada à f.d. F , $U(t) := \left(\frac{1}{1-F}\right)^{\leftarrow}(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right)$, $t \in [1, \infty[$.

Teorema 2.2.8 (Condição suficiente de von Mises). *Seja F uma f.d. absolutamente contínua e suponha-se a existência de F'' e que $f = F'$ seja positiva numa vizinhança à esquerda de x^F . Se*

$$\lim_{t \rightarrow x^F} \frac{d}{dt} \left[\frac{1 - F(t)}{f(t)} \right] = \xi, \quad (2.20)$$

ou equivalentemente,

$$\lim_{t \rightarrow x^F} \frac{(1 - F(t)) F''(t)}{f^2(t)} = -\xi - 1, \quad (2.21)$$

então

$$F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi}).$$

Observação 2.2.4. *Sob a validade de (2.21) tem-se $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$ com*

$$b_n = F^{\leftarrow}\left(1 - \frac{1}{n}\right) = U(n), \quad a_n = \frac{1}{nf(b_n)} = nU'(n).$$

Observação 2.2.5. *Note-se que este teorema não só indica as constantes de atração como também o valor do parâmetro de forma da distribuição EV. Contudo esta condição é apenas suficiente e não necessária e apenas se aplica a distribuições absolutamente contínuas.*

O Teorema 2.2.8 no caso $\xi = 0$ pode-se escrever:

$$\text{se } \lim_{t \rightarrow x^F} \frac{d}{dt} \left[\frac{1 - F(t)}{f(t)} \right] = 0 \left(\Leftrightarrow \lim_{t \rightarrow x^F} \frac{(1 - F(t)) F''(t)}{f^2(t)} = -1 \right), \quad (2.22)$$

então $F \in \mathcal{D}_{\mathcal{M}}(EV_0)$, ou seja, $F \in \mathcal{D}_{\mathcal{M}}(\Lambda)$.

Para $\xi \neq 0$, são possíveis condições mais simples como as que são apresentadas no teorema seguinte.

Teorema 2.2.9. *Seja F uma f.d. absolutamente contínua com densidade $f = F'$.*

i) *Se $x^F = +\infty$, e para algum $\xi > 0$,*

$$\lim_{t \rightarrow \infty} \frac{tf(t)}{1 - F(t)} = \frac{1}{\xi} = \alpha, \quad (2.23)$$

então $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$ com $\xi > 0$, ou seja, $F \in \mathcal{D}_{\mathcal{M}}(\Phi_{\alpha})$ com $\alpha = \frac{1}{\xi} > 0$;

ii) *se $x^F < \infty$, e para algum $\xi < 0$,*

$$\lim_{t \rightarrow x^F} \frac{(x^F - t)f(t)}{1 - F(t)} = -\frac{1}{\xi} = \alpha, \quad (2.24)$$

então $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$, com $\xi < 0$, ou seja, $F \in \mathcal{D}_{\mathcal{M}}(\Psi_{\alpha})$ com $\alpha = -\frac{1}{\xi} > 0$.

A condição von Mises (Teorema 2.2.8) pode ser apresentada numa nova versão, da autoria de de Haan, recorrendo à função quantil recíproca de F , $U(\cdot)$.

Corolário 2.2.1 (Condição suficiente de von Mises, recorrendo a U). *Sob as condições do Teorema 2.2.8, se $\lim_{t \rightarrow \infty} \frac{tU''(t)}{U'(t)} = \xi - 1$, o que é equivalente a (2.20), então, $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$.*

As condições referidas atrás, quando particularizadas a cada um dos domínios Gumbel, Fréchet e Weibull podem também ser reescritas.

Para $\xi = 0$ tem-se, aplicando diretamente o Teorema anterior que, se

$$\lim_{t \rightarrow \infty} \frac{tU''(t)}{U'(t)} = -1$$

então $F \in \mathcal{D}_{\mathcal{M}}(EV_0)$, ou seja, $F \in \mathcal{D}_{\mathcal{M}}(\Lambda)$.

Para $\xi \neq 0$ veja-se o corolário seguinte.

Corolário 2.2.2. *Sob as condições do Teorema 2.2.9,*

1. *se*

$$\lim_{t \rightarrow \infty} \frac{tU'(t)}{U(t)} = \xi,$$

então $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$ com $\xi > 0$;

2. se

$$\lim_{t \rightarrow \infty} \frac{tU'(t)}{U(\infty) - U(t)} = -\xi,$$

então $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$, com $\xi < 0$.

No caso de não ser possível garantir a existência da segunda derivada de F , o teorema que se segue, proposto por de Haan (1970), fornece uma condição suficiente para o domínio de atração da Gumbel.

Teorema 2.2.10 (de Haan, 1970). *Seja F uma f.d. absolutamente contínua no intervalo $]x_0, x^F]$, com $x_0 < x^F$ e com f.d.p. positiva f . Se*

$$\lim_{t \rightarrow x^F} \frac{f(t) \int_t^{x^F} (1 - F(t)) dt}{(1 - F(t))^2} = 1, \quad (2.25)$$

então $F \in \mathcal{D}_{\mathcal{M}}(\Lambda)$.

Gnedenko (1943) estabeleceu um conjunto de condições necessárias e suficientes para que uma f.d. F pertença ao domínio de atração para máximos da distribuição EV_{ξ} com $\xi \in \mathbb{R}$. A completa caracterização deve-se mais tarde a de Haan (1970).

Teorema 2.2.11. *Condições necessárias e suficientes:*

i) $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$ com $\xi > 0$ $\left(F \in \mathcal{D}_{\mathcal{M}}(\Phi_{\alpha}), \alpha = \frac{1}{\xi} > 0 \right)$ se e só se

$$x^F = \infty \quad e \quad \lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\frac{1}{\xi}} = -\ln(\Phi_{\alpha}(x)) = x^{-\alpha}, \quad \forall x > 0. \quad (2.26)$$

Neste caso pode-se escolher $b_n = 0$ e $a_n = F^{\leftarrow} \left(1 - \frac{1}{n} \right) = U(n)$.

ii) $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$ com $\xi < 0$ $\left(F \in \mathcal{D}_{\mathcal{M}}(\Psi_{\alpha}), \alpha = -\frac{1}{\xi} > 0 \right)$ se e só se

$$x^F < \infty \quad e \quad \lim_{t \rightarrow 0^+} \frac{1 - F(x^F - tx)}{1 - F(x^F - t)} = x^{-\frac{1}{\xi}} = -\ln(\Psi_{\alpha}(-x)) = x^{\alpha}, \quad \forall x > 0. \quad (2.27)$$

Pode então escolher-se $b_n = x^F$ e $a_n = x^F - F^{\leftarrow} \left(1 - \frac{1}{n} \right) = x^F - U(n)$.

iii) $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$ com $\xi = 0$ ($F \in \mathcal{D}_{\mathcal{M}}(\Lambda)$) se e só se existir uma função positiva $e(\cdot)$

tal que,

$$x^F \leq \infty \quad e \quad \lim_{t \rightarrow x^F} \frac{1 - F(t + xe(t))}{1 - F(t)} = e^{-x} = -\ln(\Lambda(x)), \quad \forall x \in \mathbb{R}, \quad (2.28)$$

$$com \int_t^{x^F} (1 - F(s)) ds < \infty.$$

A função $e(\cdot)$ pode ser a função do excesso médio dada por,

$$e(t) := \mathbb{E}[X - t | X > t] = \frac{\int_t^{x^F} (1 - F(s)) ds}{1 - F(t)}, \quad \forall t < x^F. \quad (2.29)$$

As constantes normalizadoras podem ser dadas por $b_n = F^{\leftarrow}\left(1 - \frac{1}{n}\right) = U(n)$ e $a_n = \mathbb{E}[X - b_n | X > b_n]$.

2.2.3 Variação Regular e caracterização dos domínios de atração

A teoria de variação regular, desenvolvida em 1930 por Karamata, foi aplicada por de Haan (1970) na caracterização dos domínios de atração em EVT. Os resultados apresentados de seguida poderão ser consultados em de Haan (1970, 1976), de Haan (1984) e em de Haan and Ferreira (2006).

Definição 2.2.3 (Função de Variação Regular). *Uma função mensurável $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ é de variação regular em ∞ com índice de variação $\alpha \in \mathbb{R}$, e escreve-se $f \in RV_\alpha$, se para qualquer $x > 0$,*

$$\lim_{t \rightarrow \infty} \frac{f(tx)}{f(t)} = x^\alpha. \quad (2.30)$$

Se $\alpha = 0$ diz-se que f é de variação lenta, sendo usualmente denotada por $L(\cdot)$. As funções de variação regular com índice α podem ser decompostas num produto de duas funções, i.e., $f(x) = x^\alpha L(x)$ com $L \in RV_0$.

Observação 2.2.6. A condição **i)** do Teorema 2.2.11 pode-se escrever na seguinte forma em termos da cauda direita, $\bar{F} = 1 - F$,

$$F \in \mathcal{D}_{\mathcal{M}}(EV_\xi) \text{ com } \xi > 0 \quad \text{sse} \quad \bar{F} \in RV_{-\frac{1}{\xi}}.$$

Para a caracterização dos domínios de atração para máximos das famílias de dis-

tribuições Fréchet e Weibull é suficiente a teoria das funções de variação regular. No entanto, para se conseguir a caracterização do domínio de atração da Gumbel é necessário definir algumas extensões da classe das funções de variação regular, nomeadamente as funções de Π variação e variação regular generalizada.

Definição 2.2.4 (Variação Regular Generalizada). *Uma função $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ mensurável diz-se ser de variação regular generalizada se existir uma função positiva $a : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ tal que para qualquer $x > 0$ e para algum $\alpha \in \mathbb{R}$,*

$$\lim_{t \rightarrow \infty} \frac{f(tx) - f(t)}{a(t)} = \begin{cases} \frac{x^\alpha - 1}{\alpha}, & \text{se } \alpha \neq 0 \\ \ln x & \text{se } \alpha = 0, \end{cases}$$

e escreve-se $f \in ERV_\alpha$. A função a é denominada de função auxiliar para f .

Observação 2.2.7 (Π -Variação). *Uma função $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ mensurável diz-se ser de Π -variação, se existir uma função positiva $a : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ tal que para qualquer $x > 0$,*

$$\lim_{t \rightarrow \infty} \frac{f(tx) - f(t)}{a(t)} = \ln(x),$$

i.e. $f \in ERV_0$ e escreve-se $f \in \Pi(a)$.

Recorrendo à teoria das funções de variação regular, de variação regular generalizada e de Π -variação, juntamente com a função quantil recíproca de F , $U(\cdot)$, de Haan determinou novas condições para que uma f.d. pertença a um determinado domínio de atração e reformulou as condições de von Mises e de Gnedenko.

Teorema 2.2.12 (de Haan and Ferreira, 2006, Theorem 1.1.6). *Seja EV_ξ a f.d. de valores extremos apresentada em (2.17), com $\xi \in \mathbb{R}$. As seguintes afirmações são equivalentes:*

1. *Existem sucessões $\{a_n\}_{n \geq 1} \in \mathbb{R}^+$ e $\{b_n\}_{n \geq 1} \in \mathbb{R}$, tais que*

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = EV_\xi(x) \Leftrightarrow F \in \mathcal{D}_M(EV_\xi),$$

para todo o x tal que $1 + \xi x > 0$.

2. Existe uma função positiva $a(\cdot)$ tal que,

$$\lim_{t \rightarrow \infty} t(1 - F(a(t)x + U(t))) = -\ln EV_\xi(x) = (1 + \xi x)^{-\frac{1}{\xi}},$$

para todo o x tal que $1 + \xi x > 0$.

3. Existe uma função positiva $a(\cdot)$ tal que, para $x > 0$, se tem,

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = \begin{cases} \frac{x^\xi - 1}{\xi}, & \xi \neq 0 \\ \ln x, & \xi = 0. \end{cases} \quad (2.31)$$

4. Existe uma função positiva $e(\cdot)$ tal que,

$$\lim_{t \rightarrow x^F} \frac{\overline{F}(t + xe(t))}{\overline{F}(t)} = (1 + \xi x)^{-\frac{1}{\xi}}$$

com $1 + \xi x > 0$.

Tem-se também que $b_n := U(n)$, $a_n := a(n)$ e 4. verifica-se com $e(t) = a\left(\frac{1}{\overline{F}(t)}\right)$.

Observação 2.2.8. Relativamente à condição 4. do teorema anterior, de Haan demonstrou que $e(\cdot)$ depende do sinal de ξ e que pode ser expressa como,

$$e(t) = \begin{cases} \xi t, & \xi > 0, \\ -\xi (x^F - t), & \xi < 0, \\ \frac{\int_t^{x^F} (1 - F(s)) ds}{1 - F(t)}, & \xi = 0. \end{cases} \quad (2.32)$$

No caso $\xi = 0$, $e(\cdot)$ é a função do excesso médio já referida anteriormente, em (2.29).

Observação 2.2.9. A igualdade em (2.31) é equivalente a afirmar que a função $U(\cdot)$ é de variação regular generalizada. Sendo assim, $U \in ERV_\xi$ é uma condição necessária e suficiente para se ter $F \in \mathcal{D}_M(EV_\xi)$. Esta condição é referida como a condição de primeira ordem, i.e.,

$$F \in \mathcal{D}_M(EV_\xi) \Leftrightarrow U \in ERV_\xi. \quad (2.33)$$

Veja-se também, recorrendo à função U , um lema que estabelece condições necessárias para que $F \in \mathcal{D}_M(EV_\xi)$.

Lema 2.2.1 (Condições necessárias para $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$). *Admita-se que $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$, com $\xi \in \mathbb{R}$.*

1. *Se $\xi > 0$, então*

$$U(\infty) = \infty \quad e \quad \lim_{t \rightarrow \infty} \frac{U(t)}{a(t)} = \frac{1}{\xi}.$$

2. *Se $\xi < 0$, então*

$$U(\infty) < \infty \quad e \quad \lim_{t \rightarrow \infty} \frac{U(\infty) - U(t)}{a(t)} = -\frac{1}{\xi}.$$

3. *Se $\xi = 0$, então*

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = 1 \Leftrightarrow U \in RV_0, \forall_{x>0} \quad e \quad \lim_{t \rightarrow \infty} \frac{a(t)}{U(t)} = 0, \quad se \quad U(\infty) = \infty$$

e

$$\lim_{t \rightarrow \infty} \frac{U(\infty) - U(tx)}{U(\infty) - U(t)} = 1 \Leftrightarrow U(\infty) - U \in ERV_0, \forall_{x>0}$$

$$e \quad \lim_{t \rightarrow \infty} \frac{a(t)}{U(\infty) - U(t)} = 0, \quad se \quad U(\infty) < \infty.$$

Tem-se ainda que,

$$\lim_{t \rightarrow \infty} \frac{a(tx)}{a(t)} = 1, \forall_{x>0} \quad i.e., \quad a \in RV_0.$$

Recorrendo à teoria das funções de variação regular, juntamente com a função quantil recíproca de F , $U(\cdot)$, pode-se reescrever as condições necessárias e suficientes do Teorema de Gnedenko (Teorema 2.2.11).

Teorema 2.2.13 (Condições necessárias e suficientes para $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$). *Pode-se garantir as condições seguintes:*

1. $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$, com $\xi > 0$, se e só se

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^{\xi} \Leftrightarrow U \in RV_{\xi}, \forall_{x>0}. \quad (2.34)$$

2. $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$, com $\xi < 0$, se e só se

$$U(\infty) < \infty \quad e \quad \lim_{t \rightarrow \infty} \frac{U(\infty) - U(tx)}{U(\infty) - U(t)} = \lim_{t \rightarrow \infty} \frac{x^F - U(tx)}{x^F - U(t)} = x^{\xi} \Leftrightarrow x^F - U \in RV_{\xi}, \quad \forall x > 0.$$

3. $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$, com $\xi = 0$, se e só se

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = \ln x \Leftrightarrow U \in \Pi(a).$$

2.2.4 Condições de primeira e segunda ordem

Como já se referiu na secção anterior, a **condição de primeira ordem** estabelece que $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$ se e só se existir uma função positiva $a(\cdot)$ tal que para todo o $x > 0$,

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = \begin{cases} \frac{x^{\xi} - 1}{\xi}, & \xi \neq 0 \\ \ln x, & \xi = 0, \end{cases}$$

i.e., se se tiver a validade de (2.31).

No caso de modelos de cauda pesada, escolhendo $a(t) = \xi U(t)$, a condição de primeira ordem pode ser escrita em termos da função quantil recíproca da f.d. F , $U(\cdot)$, como se pode ver em (2.34). De forma mais específica,

$$F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi > 0}) \Leftrightarrow U \in RV_{\xi} \Leftrightarrow \lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^{\xi} \Leftrightarrow \lim_{t \rightarrow \infty} \frac{\ln U(tx) - \ln U(t)}{\ln x} = \xi, \quad (2.35)$$

ou, equivalentemente, em termos da cauda direita da f.d. F , $\bar{F}(\cdot)$,

$$F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi > 0}) \Leftrightarrow \bar{F} \in RV_{-\frac{1}{\xi}} \Leftrightarrow \lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\frac{1}{\xi}}. \quad (2.36)$$

A condição de primeira ordem pode ser utilizada para garantir a consistência dos estimadores semi-paramétricos de parâmetros de acontecimentos extremos, como se irá estudar nos capítulos posteriores.

Para se conhecer o comportamento assintótico, em distribuição, desses estimadores é ainda usual considerar uma condição de variação regular de segunda ordem rela-

tiva ao comportamento de segunda ordem da cauda do modelo subjacente à amostra. Esta condição permite medir a velocidade de convergência da condição de primeira ordem (2.31), i.e., a velocidade de convergência de $\frac{U(tx) - U(t)}{a(t)}$ para $\frac{x^\xi - 1}{\xi}$, ou para $\ln x$, dependendo de um parâmetro de segunda ordem, $\rho \leq 0$, que dessa forma controla a velocidade de convergência de valores máximos, linearmente normalizados, para a lei limite em (2.17). Quanto maior for o valor de $|\rho|$ maior será essa velocidade.

A condição de segunda ordem pode ser escrita da seguinte forma:

$$\lim_{t \rightarrow \infty} \frac{\frac{U(tx) - U(t)}{a(t)} - \frac{x^\xi - 1}{\xi}}{\mathcal{A}(t)} = H(x, \xi, \rho), \quad \forall x > 0, \rho \leq 0, \quad (2.37)$$

onde a função $\mathcal{A}(t)$ converge para 0, quando $t \rightarrow \infty$, e a função $H(x, \xi, \rho)$ não é múltipla de $\frac{x^\xi - 1}{\xi}$.

Para determinadas funções $a(\cdot)$ e $\mathcal{A}(\cdot)$, a função H pode assumir a forma,

$$H(x, \xi, \rho) = \frac{1}{\rho} \left(\frac{x^{\xi+\rho} - 1}{\xi + \rho} - \frac{x^\xi - 1}{\xi} \right), \quad x > 0, \quad \rho \leq 0.$$

Se $\rho < 0$, H pode assumir a seguinte expressão,

$$H(x, \xi, \rho) = \frac{1}{\rho} \left(\frac{x^{\xi+\rho} - 1}{\xi + \rho} \right), \quad x > 0.$$

A condição de segunda ordem referida em (2.37) caracteriza a cauda de modelos no domínio de atração para máximos de qualquer uma das três distribuições limite e é complexa (para detalhes, veja-se de Haan and Ferreira, 2006). No entanto se F for um modelo de caudas pesadas é possível substituir esta condição por outra equivalente e mais simples. Neste caso, a condição de segunda ordem mede a velocidade de convergência de $\frac{1 - F(tx)}{1 - F(t)}$ para $x^{-\frac{1}{\xi}}$, ou de $\frac{U(tx)}{U(t)}$ para x^ξ , através do parâmetro de segunda ordem $\rho \leq 0$ e pode ser escrita da seguinte forma,

$$\lim_{t \rightarrow \infty} \frac{\frac{1 - F(tx)}{1 - F(t)} - x^{-\frac{1}{\xi}}}{\bar{A}(t)} = \begin{cases} x^{-\frac{1}{\xi}} \frac{x^{\frac{\rho}{\xi}} - 1}{\frac{\rho}{\xi}}, & \text{se } \rho \neq 0 \\ x^{-\frac{1}{\xi}} \ln x, & \text{se } \rho = 0, \end{cases} \quad (2.38)$$

para qualquer $x > 0$, onde a função $\bar{A}(\cdot)$ é de sinal constante, verificando $|\bar{A}| \in RV_{\frac{\rho}{\xi}}$.

Em termos de U , pode ser escrita,

$$\lim_{t \rightarrow \infty} \frac{\frac{U(tx)}{U(t)} - x^\xi}{A(t)} = \begin{cases} x^\xi \frac{x^\rho - 1}{\rho}, & \text{se } \rho \neq 0 \\ x^\xi \ln x, & \text{se } \rho = 0, \end{cases} \quad (2.39)$$

para qualquer $x < 0$, onde a função $A(\cdot)$ verifica $|A| \in RV_\rho$.

Tem-se também que $A(t) \sim \xi^2 \overline{A}(U(t))$ (Geluk and de Haan, 1987).

Observação 2.2.10. *Um modelo F que verifique uma das condições de variação regular de uma determinada ordem, verifica também todas as outras de ordem inferior.*

A condição (2.39) permite obter uma aproximação para o quociente $\frac{U(tx)}{U(t)}$,

$$\frac{U(tx)}{U(t)} \sim x^\xi \left(1 + \left(\frac{x^\rho - 1}{\rho} \right) A(t) \right),$$

e recorrendo ao desenvolvimento assintótico de ordem um da função $\ln(1+x)$, tem-se

$$\ln \frac{U(tx)}{U(t)} - \xi \ln x \sim \left(\frac{x^\rho - 1}{\rho} \right) A(t).$$

Desta forma a condição de segunda ordem (2.39) é assintoticamente equivalente a

$$\lim_{t \rightarrow \infty} \frac{\ln U(tx) - \ln U(t) - \xi \ln x}{A(t)} = \begin{cases} (x^\rho - 1)/\rho, & \text{se } \rho < 0, \\ \ln x, & \text{se } \rho = 0, \end{cases} \quad (2.40)$$

$\forall x > 0$, sendo muitas vezes utilizada nesta forma.

Admitindo que se lida com a classe de modelos Hall-Welsh (Hall and Welsh, 1985), com uma função de cauda direita,

$$\overline{F}(x) = (x/C)^{-1/\xi} \left(1 + \beta(x/C)^{\rho/\xi} / \rho + o(x^{\rho/\xi}) \right), \quad \text{quando } x \rightarrow \infty,$$

$C > 0$, $\beta \neq 0$ e $\rho < 0$, ou, de forma equivalente, com uma função quantil recíproca,

$$U(t) = C t^\xi \left(1 + \xi \beta t^\rho / \rho + o(t^\rho) \right), \quad \text{quando } t \rightarrow \infty, \quad (2.41)$$

pode-se afirmar, com (β, ρ) o vetor de parâmetros de segunda ordem, que a condição de segunda ordem, em (2.40), se verifica com $A(t) = \xi \beta t^\rho$, $\rho < 0$. Modelos como a log-Gama, onde $\rho = 0$, são assim excluídos desta classe. A distribuição de Pareto padrão, onde $\rho = -\infty$ é também excluída. Mas a maioria dos modelos de cauda pesada utilizados em aplicações, como o modelo EV_ξ , em (2.17), a distribuição Fréchet, $\Phi_{1/\xi}$, em (2.16b), e a distribuição de Student, t_ν , entre outras, pertencem à classe Hall-Welsh. Neste trabalho será sempre admitido que $\rho < 0$.

2.2.5 Estatísticas ordinais centrais e intermédias

O quantil empírico $X_{k(n):n}$, $k(n) := \lfloor nq \rfloor + 1$, $0 < q < 1$, com $\lfloor x \rfloor$ a parte inteira de um número x , é uma e.o. central. Se F for uma f.d. absolutamente contínua, com f.d.p. f contínua e positiva em $F^{-1}(q)$ ($f(F^{-1}(q)) > 0$), a variável $X_{k(n):n}$ após normalização adequada, tem distribuição assintoticamente normal.

Teorema 2.2.14 (Distribuição assintótica de uma e.o. central). *Seja F uma f.d. absolutamente contínua, com f.d.p. f contínua e positiva em $F^{-1}(q)$, com $0 < q < 1$. Para $k(n) := \lfloor nq \rfloor + 1$,*

$$\sqrt{n}f(F^{-1}(q)) \frac{X_{k(n):n} - F^{-1}(q)}{\sqrt{q(1-q)}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1). \quad (2.42)$$

Observação 2.2.11. *O quantil empírico, $X_{k(n):n}$ é um estimador consistente de $F^{-1}(q)$.*

Relativamente à distribuição limite de uma e.o. intermédia verifica-se que depende da velocidade de crescimento de $k = k(n)$ (Falk, 1989).

Teorema 2.2.15 (Distribuição assintótica de uma e.o. intermédia). *Seja F uma f.d. absolutamente contínua, satisfazendo uma das condições de von Mises, apresentadas nos Teoremas 2.2.9 ou 2.2.8, e $k = k(n)$ uma sucessão intermédia verificando (2.13). Então existem constantes normalizadoras $\{a_n\} \in \mathbb{R}^+$ e $\{b_n\} \in \mathbb{R}$, tais que,*

$$\frac{X_{n-k+1:n} - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1). \quad (2.43)$$

As constantes normalizadoras podem ser dadas por: $a_n = \sqrt{k}/[nf(b_n)]$ e $b_n = F^{\leftarrow}(1 - k/n)$.

Corolário 2.2.3. *Seja Y uma v.a. com distribuição de Pareto unitária e $k = k(n)$ uma sucessão intermédia, definida em (2.13). Então, escolhendo $a_n = n/(k\sqrt{k})$, $b_n = n/k$, tem-se*

$$\sqrt{k} \left(\frac{k}{n} Y_{n-k+1:n} - 1 \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1). \quad (2.44)$$

Relativamente ao modelo Pareto apresentam-se ainda mais duas propriedades que irão ser referenciadas ao longo deste trabalho.

Proposição 2.2.1. *Seja Y uma v.a. com distribuição de Pareto unitária e $k = k(n)$ uma sucessão intermédia então,*

$$\frac{k}{n} Y_{n-k:n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 1. \quad (2.45)$$

Demonstração. O resultado é verificado atendendo a que $\mathbb{E} \left[\frac{k}{n} Y_{n-k:n} \right] = 1$ e $\mathbb{V}ar \left[\frac{k}{n} Y_{n-k:n} \right] = \frac{n-k}{n(k-1)} \xrightarrow[n \rightarrow \infty]{} 0$, pois $k = k(n)$ é uma sucessão intermédia. \square

Proposição 2.2.2. *Nas condições da proposição anterior, e $\forall p \in \mathbb{R}$, é válida a seguinte representação em distribuição,*

$$\left(\frac{k}{n} Y_{n-k:n} \right)^p \stackrel{d}{=} 1 + \frac{p}{\sqrt{k}} Z_k (1 + o_{\mathbb{P}}(1)), \quad (2.46)$$

onde $Z_k = \sqrt{k} \left(\frac{k}{n} Y_{n-k:n} - 1 \right)$ é uma v.a. com distribuição assintótica normal padrão.

Demonstração. O resultado obtém-se aplicando o desenvolvimento em série de Taylor de primeira ordem da função x^p no ponto $x = 1$. \square

2.3 Inferência estatística em EVT

Para além da estimação dos parâmetros de forma, escala e localização dos modelos já referidos na Secção 2.2, é também de grande importância em EVT a estimação de quantis elevados de probabilidade $1 - q$, com q pequeno. Estes quantis são níveis elevados, usualmente designados por *níveis de retorno*, associados aos *períodos de retorno*, $1/q$, i.e., o tempo médio esperado entre excedências independentes de um nível elevado específico. É também de grande importância, a estimação da probabilidade de excedência de um determinado nível elevado ou o limite superior (ou inferior)

de suporte $x^F := \sup\{x : F(x) < 1\}$ (ou $x_F := \inf\{x : F(x) > 0\}$). Contudo apenas a estimação de EVI é objeto de estudo deste trabalho.

A inferência estatística em EVT deve ser baseada em observações extremas. Contudo existem várias formas de definir tais observações o que leva a diferentes abordagens de estimação desses parâmetros. Os trabalhos de Gomes *et al.* (2008c), Beirlant *et al.* (2012), Gomes and Guillou (2015) fornecem uma revisão geral das abordagens e métodos de estimação que têm sido desenvolvidos no passado e sugerem alguns temas que necessitam ainda de ser aprofundados e estudados.

A propriedade de max-estabilidade da f.d. EV_ξ , que estabelece que o máximo de uma a.a. de dimensão $m \times l$, para um valor grande de l , tem distribuição do mesmo tipo que o máximo de uma sucessão de m máximos, cada um dos quais máximos de l observações, sugeriu o aparecimento da primeira abordagem para modelar dados extremos, o chamado método dos *Máximos Anuais*, *Máximos de Blocos* (BM, do inglês "*Block Maxima*") ou método de *Gumbel* (Gumbel, 1958).

Nessa abordagem, a amostra de dimensão n é dividida em m sub-amostras (usualmente m corresponde ao número de anos observados) de dimensão l ($n = m \times l$). O modelo EV_ξ ou um dos modelos Gumbel, Fréchet ou Weibull, com os parâmetros $\xi \in \mathbb{R}$, $\lambda \in \mathbb{R}$ e $\delta \in \mathbb{R}^+$ desconhecidos, são então ajustados aos m valores máximos das correspondentes m sub-amostras.

Contudo em muitas aplicações não existe uma forma natural de definir os blocos de observações. Também pode ocorrer que o máximo num bloco tenha um valor inferior a alguns valores noutra bloco, o que implica que, na abordagem acima referida, esses valores extremos não irão ser incluídos nos dados a analisar. Por conseguinte, a metodologia BM pode não ser o método mais adequado para estudar o comportamento de valores extremos.

Uma outra metodologia consiste em fixar um nível elevado ou limiar, u , e definir como valores extremos todas as observações acima desse valor. A ideia é assim

ajustar o modelo referido em (2.18) aos excessos acima do nível elevado, u . Este método introduzido por Smith (1984) e conhecido como método POT (do inglês, "*Peaks Over Threshold*"), usa informação relevante que pode ser perdida pelo método BM.

Outra abordagem, de alguma forma paralela à anterior, consiste em considerar as k e.o. superiores da amostra. Nesta metodologia, geralmente denotada k -LO (do inglês, "*k-Largest Observations*"), quando a dimensão da amostra n é elevada e k um valor fixo, a inferência é baseada na estrutura multivariada das k estatísticas de ordem superior, referida em (2.19). Este modelo foi desenvolvido e estudado por Weissman (1978) e Gomes (1981), entre outros.

Note-se que a utilização do método POT necessita da escolha de um limiar adequado, u , o que é semelhante à escolha do número, k , de estatísticas de ordem superior a serem tidas em consideração na abordagem k -LO.

Também se pode pensar em combinar as abordagens BM e a k -LO. Em cada uma das m sub-amostras, pode considerar-se algumas estatísticas de ordem superior, e, neste caso, a inferência é baseada nos m vetores aleatórios de dimensão k . Estes m vetores, depois de terem sido devidamente normalizados pelas constantes, $\{a_n\} \in \mathbb{R}^+$ e $\{b_n\} \in \mathbb{R}$, são bem modelados pelo modelo EV_ξ -multivariado definido em (2.19). Esta metodologia é conhecida por abordagem EV_ξ -multidimensional.

A abordagem paramétrica foi a primeira a surgir tendo em consideração as distribuições limite referidas na Secção 2.2.

2.3.1 Abordagem paramétrica

Para estimar os diversos parâmetros da lei de valores extremos, vários procedimentos têm sido propostos: (i) métodos gráficos; (ii) métodos baseados nos momentos e (iii) métodos de verosimilhança. Todos estes procedimentos têm sido extensivamente estudados e aplicados na abordagem paramétrica clássica.

No contexto da abordagem BM, o método de estimação de máxima verosimilhança

(ML, do inglês "*Maximum Likelihood*"), pode revelar algumas dificuldades com as denominadas condições de regularidade. Smith (1985) mostrou que a propriedade usual de normalidade assintótica verifica-se desde $\xi > -0.5$. Recentemente, Zhou (2009, 2010) mostrou que os estimadores ML verificam a propriedade de normalidade assintótica para $\xi > -1$. Esta condição, que não é verificada para as distribuições de cauda muito leve, é satisfeita para grande parte das aplicações.

A normalidade assintótica, que permite obter intervalos de confiança, não é muito precisa, pois a aproximação da verdadeira distribuição amostral do estimador à normal é bastante pobre. Uma alternativa, e um método de estimação geralmente mais exato, baseia-se na função verosimilhança de perfil. Dado um vetor de parâmetros θ , a função *log-verosimilhança de perfil* da componente θ_i é definida como $\log L_p(\theta_i) := \max_{\theta_{-i}} \log L(\theta_i, \theta_{-i})$, onde θ_{-i} denota um vetor com todas as componentes de vetor θ excluindo θ_i . Para cada valor de θ_i , a log-verosimilhança de perfil é definida como o máximo da função log-verosimilhança em relação às outras componentes do vetor.

Assim, por exemplo, para a estimação de ξ no modelo *EV*,

$$\log L_p(\xi) := \max_{\lambda, \delta | \xi} \log L(\lambda, \delta, \xi).$$

Sob condições adequadas de regularidade (ver Beirlant *et al.*, 2004), para n suficientemente elevado, tem-se que a função,

$$D_p(\xi) := 2\{\log L(\hat{\lambda}, \hat{\delta}, \hat{\xi}) - \log L_p(\xi)\} \sim \chi^2_{(1)},$$

onde $\hat{\lambda}$, $\hat{\delta}$ e $\hat{\xi}$ são os estimadores de máxima verosimilhança de λ , δ e ξ , respectivamente. Esta propriedade é usada para obter os intervalos de confiança a $100 \times (1 - \alpha)\%$ para os parâmetros da distribuição subjacente. Particularmente, para um componente singular, por exemplo ξ , o intervalo de confiança a $(1 - \alpha) \times 100\%$ é $\{\xi : D_p(\xi) \leq q_{1-\alpha}\} = \{\xi : \log L_p(\xi) \geq \log L(\hat{\lambda}, \hat{\delta}, \hat{\xi}) - \frac{q_{1-\alpha}}{2}\}$, onde $q_{1-\alpha}$ é o quantil de probabilidade $(1 - \alpha)$ de $\chi^2_{(1)}$. Portanto, a estatística da razão de log-verosimilhança de perfil

$$-2 \log \Lambda = -2 \log \left\{ \frac{L_p(\xi_0)}{L_p(\hat{\xi})} \right\} = 2\{\log L_p(\hat{\xi}) - \log L_p(\xi_0)\},$$

para testar $H_0 : \xi = \xi_0$ versus $H_1 : \xi \neq \xi_0$, tem sob a validade de H_0 , distribuição assintótica, $\chi^2_{(1)}$, quando $n \rightarrow \infty$. H_0 é rejeitada ao nível de significância α se $-2 \log \Lambda > q_{1-\alpha}$ (ver Coles (2001) e Beirlant *et al.*, (2004) para mais detalhes).

Como as estimativas ML podem ser numericamente difíceis de obter, vários métodos alternativos têm sido propostos, podendo referir-se o método dos momentos ponderados de probabilidade (PWM, do inglês "*Probability-Weighted Moments*") introduzido por Greenwood *et al.* (1979). Os momentos ponderados de probabilidade de uma v.a X , com f.d. F são definidos como,

$$M_{p,r,s} := E\{X^p [F(X)]^r [1 - F(X)]^s\}, \quad p, r, s \in \mathbb{R}.$$

Para a f.d. EV , estes momentos foram extensivamente estudados por Hosking *et al.* (1985). Considerando uma a.a. (X_1, \dots, X_m) de uma população EV , o estimador PWM, $(\hat{\lambda}, \hat{\delta})$, quando $\xi = 0$, é a solução do sistema de equações:

$$\begin{cases} \widehat{M}_{1,0,0} = \lambda + \delta \Gamma'(1) \\ 2\widehat{M}_{1,1,0} - \widehat{M}_{1,0,0} = \log 2\delta \end{cases} \quad \text{onde} \quad \widehat{M}_{1,r,0} = \frac{1}{m} \sum_{i=1}^m \left(\prod_{l=1}^r \frac{i-l}{m-l} \right) X_{i:m},$$

com $X_{1:m} \leq X_{2:m} \leq \dots \leq X_{m:m}$ as estatísticas ordinais ascendentes associadas à a.a. (X_1, X_2, \dots, X_m) .

Para $\xi < 1$ e $\xi \neq 0$, pode obter-se o estimador PWM, $(\hat{\lambda}, \hat{\delta}, \hat{\xi})$, resolvendo o sistema de equações,

$$\begin{cases} \widehat{M}_{1,0,0} = \lambda - \frac{\delta}{\xi} (1 - \Gamma(1 - \xi)) \\ 2\widehat{M}_{1,1,0} - \widehat{M}_{1,0,0} = \frac{\delta}{\xi} \Gamma(1 - \xi) (2^\xi - 1) \\ \frac{3\widehat{M}_{1,2,0} - \widehat{M}_{1,0,0}}{2\widehat{M}_{1,1,0} - \widehat{M}_{1,0,0}} = \frac{3^\xi - 1}{2^\xi - 1} \end{cases}.$$

Neste método verifica-se a normalidade assintótica para o estimador PWM (λ, δ, ξ) desde que $\xi < 0.5$ e $m \rightarrow \infty$ (ver Beirlant *et al.*, 2004).

Na abordagem POT, os estimadores ML de ξ e $\tilde{\delta}$ no modelo GP foram estudados

por Smith (1984). Detalhes deste procedimento podem também ser visto em Davison and Smith (1990), que sintetiza os trabalhos de Smith (1984) e de Davison (1984). Também neste caso, o método PWM, entre outros, é um método de estimação alternativo.

2.3.2 Enquadramento estatístico semi-paramétrico e estimação

Em finais dos anos setenta a estimação em EVT começou a ser realizada numa abordagem semi-paramétrica, graças ao impulso do trabalho de Haan (1970). Neste contexto não é necessário considerar um modelo paramétrico específico, dependente de parâmetros de localização, escala e forma, mas apenas se assume que, para um valor apropriado de $\xi \in \mathbb{R}$, a função de distribuição subjacente F pertence a $\mathcal{D}_{\mathcal{M}}(EV_{\xi})$, ou a um sub-domínio específico de $\mathcal{D}_{\mathcal{M}}(EV_{\xi})$, sendo ξ o parâmetro primordial a ser estimado. Os estimadores semi-paramétricos deste parâmetro baseiam-se nas k e.o. superiores da amostra, ou nos excessos acima de um nível aleatório elevado, $X_{n-k:n}$, que se entende como uma e.o. intermédia superior. Note-se que o método POT e os métodos k -LO e EV_{ξ} -*multidimensional* podem assim ser considerados também sob a perspetiva semi-paramétrica.

Neste âmbito têm sido propostos vários estimadores-EVI. Pode-se referir os estimadores semi-paramétricos clássicos tais como, o estimador de Hill (1975) para modelos de cauda pesada, i.e. para $\xi > 0$, e para $\xi \in \mathbb{R}$, o estimador de Pickands (1975), o estimador dos Momentos (Dekkers *et al.*, 1989), o estimador de Hill Generalizado (introduzido em Beirlant *et al.*, 1996), e estudado mais tarde em Beirlant *et al.*, 2005, e o estimador dos Momentos Mistos (Fraga Alves *et al.*, 2009), entre outros.

São também de referir os estimadores ML e recentemente denominados por estimadores PORT-ML (PORT-ML, do inglês "*peaks over random threshold-maximum likelihood*") depois de Araújo Santos *et al.* (2006). Esta classe de estimadores foi introduzida por Smith (1984) no contexto paramétrico, e é baseada nos excessos acima de um limiar determinístico elevado u , podendo ser facilmente revertida com base nos

excessos acima de um limiar aleatório elevado $X_{n-k:n}$,

$$W_{ik} := X_{n-i+1:n} - X_{n-k:n}, \quad 1 \leq i \leq k < n. \quad (2.47)$$

Condicionalmente a $X_{n-k:n}$, estes k excessos são aproximadamente distribuídos como o conjunto das k e.o. de topo associadas a uma amostra de dimensão k da f.d. GP dada em (2.18). Re-parametrizando o vetor de parâmetros $(\xi, \tilde{\delta})$ em (ξ, ζ) com $\zeta = \frac{\xi}{\tilde{\delta}} \in \mathbb{R}$, a solução das equações de máxima verosimilhança deu origem a um estimador explícito de ξ para $\xi > -1/2$ e a um implícito de ζ . Para um estudo aprofundado das propriedades assintóticas destes estimadores consulte-se o trabalho de Drees *et al.* (2004). Mais recentemente Zhou (2009, 2010) mostrou que este estimador de EVI é também válido para $\xi > -1$.

Os estimadores PWM foram inicialmente desenvolvidos no contexto paramétrico. Contudo, de Haan and Ferreira (2006) conceberam, no contexto semi-paramétrico e com base na distribuição GP , os estimadores de ξ denominados por GPPWM, consistentes para $\xi < 1$, lidando com a amostra dos excessos acima de um nível aleatório elevado, $X_{n-k:n}$. Também Caeiro and Gomes (2011) introduziram, com base no modelo de Pareto e na amostra das excedências de um nível aleatório elevado, $X_{n-k:n}$, uma classe de estimadores de EVI denominada, estimadores de Pareto dos Momentos Ponderados de Probabilidade (PPWM), também consistentes para $\xi < 1$.

Todos os estimadores atrás citados são invariantes na escala. No entanto apenas o estimador de Pickands, o ML e o GPPWM são também invariantes na localização. Nesta conformidade, Araújo Santos *et al.* (2006) introduziram os estimadores-EVI, invariantes na escala e na localização, denominados PORT, que são construídos a partir da amostra dos excessos acima de um nível aleatório $X_{n_q:n}$, $n_q := \lfloor nq \rfloor + 1$, i.e., são funções da amostra $\underline{X}_n^{(q)} := (X_{n:n} - X_{n_q:n}, \dots, X_{n_q+1:n} - X_{n_q:n})$. Se o modelo subjacente F tiver um limite inferior de suporte finito, pode-se eventualmente considerar o mínimo da amostra como sendo o nível aleatório referido. Outros estimadores-EVI para modelos com cauda tipo Pareto, invariantes na localização, podem ser encontrados em Fraga Alves (2001) e Ling *et al.* (2012), entre outros.

Como já foi referido assume-se que na abordagem semi-paramétrica F pertença a $\mathcal{D}_{\mathcal{M}}(EV_{\xi})$, para um valor específico de $\xi \in \mathbb{R}$, ou seja, é então necessário que seja garantida a condição de primeira ordem em (2.31).

É ainda importante mencionar que para garantir a consistência fraca dos estimadores semi-paramétricos supra-mencionados, para além de se admitir a condição de primeira ordem, tem-se também de lidar com uma sucessão intermédia $k \equiv k(n)$, definida em (2.13). Sob a validade da condição de segunda ordem em (2.37) é possível também garantir a convergência em distribuição para a normal. De facto, seja $\hat{\xi}^{\bullet}(k)$ qualquer um dos estimadores referidos e $A(t)$ uma função convergente para zero quando $t \rightarrow \infty$, diretamente relacionada com a velocidade de convergência na condição de primeira ordem, tem-se que,

$$\hat{\xi}^{\bullet}(k) \stackrel{d}{=} \xi + \frac{\sigma_{\bullet}}{\sqrt{k}} Z_k^{\bullet} + b_{\bullet} A(n/k) + o_p(A(n/k)), \quad (2.48)$$

onde Z_k^{\bullet} é uma v.a. normal padrão e $(b_{\bullet}, \sigma_{\bullet}) \in \mathbb{R} \times \mathbb{R}^+$. Assim, se $\sqrt{k}A(n/k) \rightarrow \lambda$ finito, quando $n \rightarrow \infty$, então $\sqrt{k} \left(\hat{\xi}^{\bullet}(k) - \xi \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(\lambda b_{\bullet}, \sigma_{\bullet}^2)$. A $b_{\bullet} A(n/k)$ e σ_{\bullet}^2/k dá-se o nome respetivamente de *viés assintótico* e *variância assintótica* de $\hat{\xi}^{\bullet}(k)$.

A maioria dos estimadores semi-paramétricos clássicos exibem a mesma dificuldade: uma variância elevada para limiares elevados $X_{n-k:n}$, i.e., para valores pequenos de k ; e viés elevado para limiares pequenos, i.e., para valores elevados de k . Consequentemente estes estimadores revelam uma região de estabilidade relativamente diminuta na correspondente trajetória amostral (gráfico que representa o valor das estimativas versus k) e um erro quadrático médio bastante pontiagudo, fazendo com que a escolha adaptativa do limiar constitua um problema difícil quando o critério de escolha é a estabilidade da trajetória amostral ou o valor de k que minimiza o MSE. Este problema tem levado os investigadores a construir novos estimadores-EVI designados de viés reduzido de segunda ordem, SORB. Esta abordagem tem sido considerada essencialmente para caudas pesadas e a ideia base é encontrar forma de eliminar a componente dominante do viés, $b_{\bullet} A(n/k)$, em (2.48). Pode-se mencionar os trabalhos pioneiros de Drees (1996), Peng (1998), Beirlant *et al.* (1999), Feuerverger *et al.* (1999) e Gomes *et al.* (2000). Nestes artigos foram desenvolvidos estimado-

res em que se conseguiu eliminar o viés assintótico mas com o custo de aumentar a variância. Mais recentemente o problema de "trade-off" entre viés e variância tem vindo a ser ultrapassado, para modelos de caudas pesadas, com a introdução dos denominados estimadores de viés-reduzido com variância-mínima, MVRB, tal como foi feito em Caeiro *et al.* (2005), Gomes *et al.* (2007a), Gomes *et al.* (2008b) e Caeiro *et al.* (2009), entre outros. Para modelos da classe de Hall-Welsh onde $A(t) = \xi\beta t^\rho$, esses estimadores, baseados numa estimação apropriada do par de parâmetros de segunda ordem, $(\beta, \rho) \in (\mathbb{R}, \mathbb{R}^-)$, têm um viés assintótico nulo, conseguindo manter uma variância igual à do estimador de Hill. Estimadores do parâmetro de segunda ordem ρ podem ser encontrados em Peng (1998), Gomes *et al.* (2002), Fraga Alves *et al.* (2003) e mais recentemente em Goegebeur *et al.* (2008, 2010), Ciuperca and Mercadier (2010), de Wet *et al.* (2012), Worms and Worms (2012), Deme *et al.* (2013), Henriques-Rodrigues *et al.* (2014) e Caeiro and Gomes (2014b, 2015). Para a estimação do parâmetro de segunda ordem β considere-se os artigos de Gomes and Martins (2002), Caeiro and Gomes (2006), Gomes *et al.* (2010) e Henriques-Rodrigues *et al.* (2015). Uma referência mais recente desses estimadores pode ser encontrada em Caeiro *et al.* (2016a). Algoritmos para a estimação do par (β, ρ) podem ser vistos em Gomes and Pestana (2007a, 2007b) e Gomes *et al.* (2008b). Literatura adicional sobre a estimação SORB pode-se encontrar nos trabalhos de Gomes and Martins (2001, 2004) e Gomes *et al.* (2004), entre outros. Para uma revisão geral da estimação SORB veja-se o Capítulo 6 de Reiss and Thomas (2007), Gomes *et al.* (2008c), Beirlant *et al.* (2012) e Gomes and Guillou (2015).

Pode-se referir ainda uma classe promissora de estimadores-EVI recentemente estudada em Brilhante *et al.* (2013) e Brilhante *et al.* (2014), constituindo uma generalização do estimador de Hill, que em vez de usar a média geométrica de $\frac{X_{n-i+1:n}}{X_{n-k:n}}$, usa a média de Hölder de ordem p destas mesmas estatísticas. Neste contexto veja-se também os trabalhos de Paulauskas and Vaičiulis (2013, 2017), Beran *et al.* (2014), Gomes *et al.* (2015a), Caeiro *et al.* (2016a), e Gomes *et al.* (2016a).

Capítulo 3

As Classes de Estimadores de Lehmer e outros Estimadores-EVI

Considere-se (X_1, \dots, X_n) uma amostra de dimensão n de v.a. i.i.d. com f.d. $F \in \mathcal{D}_{\mathcal{M}}(\text{EV}_{\xi})$ e sejam $X_{1:n} \leq \dots \leq X_{n:n}$ as correspondentes e.o. ascendentes.

Neste trabalho lida-se com f.d. do tipo Pareto, com um EVI positivo, ou equivalentemente, com modelos tais que $\bar{F}(x) = x^{-1/\xi}L(x)$, $\xi > 0$, com $L \in RV_0$ uma função de variação lenta em infinito, i.e., uma função de variação regular com índice de variação regular igual a zero. Estes modelos de cauda pesada são muito comuns numa variedade de áreas de aplicação tais como os seguros, as finanças, as ciências sociais, as telecomunicações, a bibliometria, a bioestatística, a informática, entre outras.

Para modelos tipo Pareto, o estimador-EVI semi-paramétrico mais conhecido é o estimador de Hill (H) (Hill, 1975), que é a média dos log-excessos,

$$V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \leq i \leq k \leq n-1, \quad (3.1)$$

tendo-se assim que,

$$\hat{\xi}^H(k) \equiv H(k) := \frac{1}{k} \sum_{i=1}^k V_{ik}, \quad 1 \leq k \leq n-1. \quad (3.2)$$

Do ponto de vista da função quantil recíproca (Definição 2.1.1) pode-se escrever a identidade distribucional $X_{i:n} \stackrel{d}{=} U(Y_{i:n})$ com $(X_{1:n}, \dots, X_{n:n})$ as e.o. associadas a uma

amostra aleatória com f.d. $F(\cdot)$ e $(Y_{1:n}, \dots, Y_{n:n})$ as e.o. de uma amostra aleatória com f.d. de Pareto unitária (ponto 2. do Teorema 2.1.3). Também se verifica a identidade distribucional (ponto 3. do Teorema 2.1.3) ,

$$\frac{Y_{n-i+1:n}}{Y_{n-k:n}} \stackrel{d}{=} Y_{k-i+1:k}, \quad 1 \leq i \leq k$$

e $\frac{k}{n} Y_{n-k:n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 1$ (Proposição 2.2.1), i.e., $Y_{n-k:n} \sim n/k$. Consequentemente, tem-se que

$$V_{ik} := \ln \frac{X_{n-i+1:n}}{X_{n-k:n}} \stackrel{d}{=} \ln \frac{U(Y_{n-i+1:n})}{U(Y_{n-k:n})} \stackrel{d}{=} \ln \frac{U(Y_{n-k:n} Y_{k-i+1:k})}{U(Y_{n-k:n})} \stackrel{d}{=} \ln \frac{U\left(\frac{n}{k} Y_{k-i+1:k}\right)}{U\left(\frac{n}{k}\right)}. \quad (3.3)$$

Admitindo que se verifica a condição de primeira ordem para $\xi > 0$ em (2.35) , fazendo $t = \frac{n}{k}$ e $x = Y_{k-i+1:k}$, e uma vez que $k \equiv k(n)$, $1 \leq k < n$, é uma sucessão intermédia de inteiros (ou seja $\frac{n}{k} \rightarrow \infty$) tem-se,

$$\lim_{\frac{n}{k} \rightarrow \infty} \ln \frac{U\left(\frac{n}{k} Y_{k-i+1:k}\right)}{U\left(\frac{n}{k}\right)} = \ln Y_{k-i+1:k}^\xi = \xi \ln Y_{k-i+1:k}.$$

Assim, e como $\ln Y_{i:n} \stackrel{d}{=} E_{i:n}$ (teorema 2.1.3),

$$V_{ik} \stackrel{d}{=} \xi \ln Y_{k-i+1:k} + o_{\mathbb{P}}(1) \stackrel{d}{=} \xi E_{k-i+1:k} + o_{\mathbb{P}}(1), \quad (3.4)$$

sendo E uma v.a. exponencial padrão. Desta forma pode-se escrever a seguinte aproximação,

$$V_{ik} \approx \xi \ln Y_{k-i+1:k} \stackrel{d}{=} \xi E_{k-i+1:k}, \quad 1 \leq i \leq k. \quad (3.5)$$

Os log-excessos, V_{ik} , $1 \leq i \leq k$, em (3.1), são assim aproximadamente as k e.o. de uma amostra de dimensão k com distribuição exponencial com valor médio ξ . Por conseguinte, sendo o estimador H , em (3.2), a média dos k log-excessos em (3.1), a lei fraca dos grandes números permite concluir que H é um estimador consistente de EVI.

Para além da média dos log-excessos, os momentos de ordem p dos log-excessos, i.e.

$$M_{k,n}^{(p)} := \frac{1}{k} \sum_{i=1}^k \left\{ \ln X_{n-i+1:n} - \ln X_{n-k:n} \right\}^p, \quad 1 \leq k \leq n-1, \quad p \geq 1, \quad (3.6)$$

introduzidos em Dekkers *et al.* (1989) [$M_{k,n}^{(1)} \equiv H(k)$] têm também desempenhado um

papel muito relevante na estimação de EVI, e até podem ser genericamente parametrizados em $p \in \mathbb{R} \setminus \{0\}$.

Uma generalização simples da média é também a média de Lehmer de ordem p (ver Havil, 2003, entre outros). Dado um conjunto de números positivos $\underline{a} = (a_1, \dots, a_k)$, esta média generaliza, não só, a média aritmética ($p = 1$), como também, a média harmónica ($p = 0$), sendo definida como,

$$L_p(\underline{a}) := \frac{\sum_{i=1}^k a_i^p}{\sum_{i=1}^k a_i^{p-1}}, \quad p \in \mathbb{R}.$$

Note-se ainda que $\lim_{p \rightarrow -\infty} L_p(\underline{a}) = \min_{1 \leq i \leq k} a_i$ e $\lim_{p \rightarrow +\infty} L_p(\underline{a}) = \max_{1 \leq i \leq k} a_i$.

O estimador-EVI H pode assim ser considerado como a média de Lehmer de ordem 1 dos k log-excessos $\underline{V} := (V_{ik}, 1 \leq i \leq k)$ em (3.1). Mais geralmente pode-se também pensar na média de Lehmer de ordem p destas estatísticas. Atendendo a (3.5),

$$V_{ik}^p \approx \xi^p E_{k-i+1:k}^p, \quad 1 \leq i \leq k,$$

e como $\mathbb{E}(E^p) = \Gamma(p+1)$ para qualquer real $p > -1$, a lei fraca dos grandes números permite deduzir que,

$$M_{k,n}^{(p)} := \frac{1}{k} \sum_{i=1}^k V_{ik}^p \xrightarrow[k \rightarrow \infty]{p} \Gamma(p+1) \xi^p.$$

Consequentemente a classe *Lehmer* (L) de estimadores de EVI, definida como

$$\hat{\xi}^{L_p}(k) \equiv L_p(k) := \frac{L_p(\underline{V})}{p} = \frac{1}{p} \frac{\sum_{i=1}^k V_{ik}^p}{\sum_{i=1}^k V_{ik}^{p-1}} = \frac{M_{k,n}^{(p)}}{p M_{k,n}^{(p-1)}}, \quad 1 \leq k \leq n-1 \quad (3.7)$$

$[L_1(k) \equiv H(k)]$, veja-se Penalva *et al.* (2016a), é uma classe de estimadores consistentes de ξ , para $\xi > 0$ e para todo o real $p > 0$.

Considera-se ainda a classe mais simples de estimadores RB de segunda ordem de EVI introduzida em Caeiro *et al.* (2005), consistente para $\xi > 0$. Com $(\hat{\beta}, \hat{\rho})$ um estimador adequado do vetor de parâmetros (β, ρ) , em (2.41), considera-se assim a classe de estimadores-EVI de Hill-corrigido (CH, do inglês “*corrected-Hill*”) definida e

dada por,

$$\hat{\xi}^{\text{CH}}(k) \equiv \text{CH}(k) := \hat{\xi}^{\text{H}}(k) \left(1 - \hat{\beta}(n/k)^{\hat{\rho}} / (1 - \hat{\rho}) \right), \quad 1 \leq k \leq n-1. \quad (3.8)$$

Os estimadores em (3.8) podem ser estimadores MVRB quando existe uma escolha adequada do nível k e uma apropriada estimação externa do vetor dos parâmetros de segunda ordem, (β, ρ) , i.e., o uso dos estimadores $\text{CH}(k)$ permite eliminar a componente dominante do viés do estimador-EVI de Hill, $\text{H}(k)$, mantendo a sua variância assintótica. Neste trabalho considera-se para estimador de ρ , aquele que é definido em Fraga Alves *et al.* (2003), e para estimador de β , o que é considerado em Gomes and Martins (2002). Para estimar esses parâmetros foi utilizado o algoritmo providenciado em Gomes and Pestana (2007b) e reproduzido em Brilhante *et al.* (2013), aí denominado por Algoritmo 5.1, Steps 2 a 4.

De uma forma mais geral que a classe em (3.8), irá eliminar-se a componente dominante do viés de $\text{L}_p(k)$, em (3.7), resultando nos estimadores-EVI de Lehmer RB,

$$\hat{\xi}^{\text{L}_p^{\text{RB}}}(k) \equiv \text{L}_p^{\text{RB}}(k) := \hat{\xi}^{\text{L}_p}(k) \left(1 - \hat{\beta}(n/k)^{\hat{\rho}} / (1 - \hat{\rho})^p \right), \quad 1 \leq k \leq n-1 \quad (3.9)$$

$[\text{L}_1^{\text{RB}}(k) \equiv \text{CH}(k)]$. Veja-se Gomes *et al.* (2016b).

Este capítulo aborda o comportamento assintótico da classe de estimadores-EVI L_p em (3.7) e L_p^{RB} em (3.9) como também o comportamento de alguns outros estimadores frequentemente estudados, procurando efectuar algumas comparações entre esses estimadores.

3.1 Outros estimadores-EVI em consideração

Para além do estimador de Hill e das previamente mencionadas *classes de estimadores de Lehmer* em (3.7) e (3.9), consideram-se também no presente trabalho os seguintes estimadores-EVI:

- O *estimador dos momentos* (M), consistente para $\xi \in \mathbb{R}$, estudado em Dekkers *et al.* (1989) e baseado nas estatísticas $M_{k,n}^{(1)}, M_{k,n}^{(2)}$, com $M_{k,n}^{(p)}$ definido em (3.6),

é dado por,

$$\hat{\xi}^M(k) \equiv M(k) := M_{k,n}^{(1)} + \frac{1}{2} \left\{ 1 - \left(M_{k,n}^{(2)} / (M_{k,n}^{(1)})^2 - 1 \right)^{-1} \right\}, \quad 1 \leq k \leq n-1. \quad (3.10)$$

- O *estimador generalizado de Hill* (GH), consistente para $\xi \in \mathbb{R}$ (Beirlant *et al.*, 1996, 2005), é baseado no estimador de Hill (3.2) e tem a forma funcional,

$$\hat{\xi}^{GH}(k) \equiv GH(k) := \hat{\xi}^H(k) + \frac{1}{k} \sum_{i=1}^k \left\{ \ln \hat{\xi}^H(i) - \ln \hat{\xi}^H(k) \right\}, \quad 1 \leq k \leq n-1. \quad (3.11)$$

- O *estimador dos momentos mistos* (MM), consistente para $\xi \in \mathbb{R}$ (Fraga Alves *et al.*, 2009), é definido por,

$$\hat{\xi}^{MM}(k) \equiv MM(k) := \frac{\hat{\varphi}_{k,n} - 1}{1 + 2 \min(\hat{\varphi}_{k,n} - 1, 0)}, \quad 1 \leq k \leq n-1, \quad (3.12)$$

com $\hat{\varphi}_{k,n} := \frac{M_{k,n}^{(1)} - L_{k,n}^{(1)}}{(L_{k,n}^{(1)})^2}$, $M_{k,n}^{(p)}$ dado por (3.6) e $L_{k,n}^{(p)}$ definido por,

$$L_{k,n}^{(p)} := \frac{1}{k} \sum_{i=1}^k \left(1 - \frac{X_{n-k:n}}{X_{n-i+1:n}} \right)^p, \quad p \geq 1. \quad (3.13)$$

- O *estimador de Pareto generalizado dos momentos ponderados de probabilidade* (GPPWM), consistente para $\xi < 1$ (de Haan and Ferreira, 2006), é baseado na amostra dos excessos acima de um nível elevado aleatório $X_{n-k:n}$ e definido por,

$$\hat{\xi}^{GPPWM}(k) \equiv GPPWM(k) := 1 - \frac{2\hat{a}_1^*(k)}{\hat{a}_0^*(k) - 2\hat{a}_1^*(k)}, \quad 1 \leq k \leq n-1, \quad (3.14)$$

com

$$\hat{a}_s^*(k) := \frac{1}{k} \sum_{i=1}^k \left(\frac{i-1}{k-1} \right)^s (X_{n-i+1:n} - X_{n-k:n}), \quad s = 0, 1.$$

- O *estimador de Pareto dos momentos ponderados de probabilidade* (PPWM), consistente para $\xi < 1$, introduzido em Caeiro and Gomes (2011) (ver também Caeiro *et al.*, 2014, 2016b), baseado na amostra das excedências, é definido por

$$\hat{\xi}^{PPWM}(k) \equiv PPWM(k) := 1 - \frac{\hat{a}_1(k)}{\hat{a}_0(k) - \hat{a}_1(k)}, \quad 1 \leq k \leq n-1, \quad (3.15)$$

onde

$$\hat{a}_0(k) := \frac{1}{k} \sum_{i=1}^k X_{n-i+1:n}, \quad \hat{a}_1(k) := \frac{1}{k} \sum_{i=1}^k \frac{i-1}{k-1} X_{n-i+1:n},$$

- A classe de estimadores, que poderão ser SORB, introduzidos e estudados em Gomes and Martins (2001), consistentes para $\xi > 0$ e $p > -1$,

$$\hat{\xi}^{\text{GM}_p}(k) \equiv \text{GM}_p(k) := \frac{M_{k,n}^{(p)}}{\Gamma(p+1) \left[M_{k,n}^{(1)} \right]^{p-1}}, \quad 1 \leq k \leq n-1 \quad (3.16)$$

$$[\text{GM}_1(k) \equiv \text{H}(k) \equiv \text{L}_1(k), \quad \text{GM}_2(k) \equiv \text{L}_2(k)].$$

- Os estimadores, que também poderão ser SORB, introduzidos e estudados em Caeiro and Gomes (2002a) (ver também Caeiro and Gomes, 2002b, 2014a), consistentes para $\xi > 0$ e $p > 0$, determinados por,

$$\hat{\xi}^{\text{CG}_{p,\delta}}(k) \equiv \text{CG}_{p,\delta}(k) := \frac{\Gamma(p)}{M_{k,n}^{(p-1)}} \left(\frac{M_{k,n}^{(\delta p)}}{\Gamma(\delta p + 1)} \right)^{1/\delta}, \quad 1 \leq k \leq n-1 \quad \delta > 0. \quad (3.17)$$

Note-se que esta família de estimadores, introduzida num contexto de procura de estimadores SORB, é bastante vasta e engloba alguns estimadores já estudados:

- $\text{CG}_{1,1}(k) \equiv \text{H}(k)$;
- Para $\delta = 2$ obtém-se a classe de estimadores estudada em Caeiro and Gomes (2002b), que generaliza o estimador $\text{CG}_{1,2}(k) = \sqrt{M_{k,n}^{(2)}/2}$, estudado em Gomes *et al.* (2000);
- Se $p = 1$ obtém-se a família de estimadores $\text{CG}_{1,\delta}(k) = \left(\frac{M_{k,n}^{(\delta)}}{\Gamma(\delta+1)} \right)^{1/\delta}$, estudada em Gomes and Martins (2001).
- A classe de estimadores L em (3.7), $\text{L}_p(k) \equiv \text{CG}_{p,1}(k)$. Neste caso tem-se um caso particular de (3.17) que não é RB.
- A classe de estimadores recentemente estudada em Brilhante *et al.* (2013) e Gomes and Caeiro (2014). Com base no facto de que o estimador de Hill em (3.2) é o logaritmo da média geométrica das estatísticas $U_{ik} := X_{n-i+1:n}/X_{n-k:n}$, e se

se considerar a média de Hölder de ordem p (MO_p) dessas mesmas estatísticas, obtêm-se os denominados estimadores MO_p , assim definidos:

$$\hat{\xi}^{\text{H}_p}(k) \equiv H_p(k) := \begin{cases} \left(1 - \left(\frac{1}{k} \sum_{i=1}^k U_{ik}^p\right)^{-1}\right) / p, & \text{se } 0 \neq p < 1/\xi, \\ H(k), & \text{se } p = 0, \end{cases} \quad (3.18)$$

com $1 \leq k \leq n-1$, $[H_0(k) \equiv H(k)]$.

- Os estimadores ML, mais recentemente denominados por estimadores PORT-ML em Araújo Santos *et al.* (2006). Esta classe de estimadores foi introduzida por Smith (1984) no contexto paramétrico com base nos excessos acima de um limiar determinístico elevado u , e pode ser facilmente revertida nos excessos acima de um limiar aleatório elevado $X_{n-k:n}$,

$$W_{ik} := X_{n-i+1:n} - X_{n-k:n}, \quad 1 \leq i \leq k \leq n-1. \quad (3.19)$$

Condicionalmente a $X_{n-k:n}$, estes k excessos são aproximadamente distribuídos como o conjunto das k e.o. associadas a uma amostra de dimensão k da f.d. GP dada em (2.18). Re-parametrizando o vetor de parâmetros $(\xi, \tilde{\delta})$ em (ξ, α) com $\alpha = \frac{\xi}{\tilde{\delta}} \in \mathbb{R}$, a solução das equações de máxima verosimilhança deu origem a um estimador explícito de ξ para $\xi > -1/2$, dependente de estimador implícito de α .

$$\hat{\xi}^{\text{ML}}(k) \equiv \hat{\xi}^{\text{ML}}(k, \hat{\alpha}_{\text{ML}}) \equiv \text{ML}(k) := \frac{1}{k} \sum_{i=1}^k \ln(1 + \hat{\alpha}_{\text{ML}} W_{ik}). \quad (3.20)$$

Para um estudo aprofundado das propriedades assintóticas destes estimadores consulte-se o trabalho de Drees *et al.* (2004). Mais recentemente Zhou (2009, 2010) mostrou que este estimador do EVI é também consistente para $\xi > -1$.

Observação 3.1.1. Um simples estimador heurístico de α é $1/X_{n-k:n}$. Ao se considerar $\hat{\alpha} = 1/X_{n-k:n}$ e os excessos W_{ik} , $1 \leq i \leq k$, em (3.19), $1 + \hat{\alpha} W_{ik} = X_{n-i+1:n}/X_{n-k:n}$. Então, $\hat{\xi}^{\text{ML}}(k, \hat{\alpha}) = \frac{1}{k} \sum_{i=1}^k \{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}$ é a média dos log-excessos V_{ik} , $1 \leq i \leq k$, i.e., obtém-se o clássico estimador H em (3.2).

Observação 3.1.2. Note-se que os estimadores GPPWM, em (3.14) e os estimadores

ML, em (3.20), são invariantes na localização e na escala. Os restantes apenas são invariantes na escala, contudo podem passar a ser invariantes na localização se se aplicar a já anteriormente referida metodologia PORT, baseando-os, não na amostra original, mas nos excessos acima de um quantil central empírico e até mesmo acima do mínimo da amostra sempre que seja possível, i.e., quando a distribuição subjacente tiver um limite inferior de suporte finito (para detalhes sobre o tema consultar, entre outros, Araújo Santos *et al.*, 2006, Gomes *et al.*, 2008a, e mais recentemente, Gomes and Henriques-Rodrigues, 2016).

3.2 Comportamento assintótico dos estimadores do EVI

Na área estatística em EVT e sempre que se trabalha com valores extremos, um modelo F é usualmente dito ser de cauda pesada (ver Secção 2.2.4), sempre que a função da cauda direita \bar{F} , em (2.14), é uma função de variação regular com índice de variação regular negativo igual a $-1/\xi$, $\xi > 0$, ou equivalentemente, se a função quantil recíproca U , em (2.8), for de variação regular com índice ξ , i.e.,

$$F \in \mathcal{D}_{\mathcal{M}}^+ := \mathcal{D}_{\mathcal{M}}(EV_{\xi})_{\xi>0} \iff \bar{F} \in RV_{-1/\xi} \iff U \in RV_{\xi}. \quad (3.21)$$

O parâmetro de segunda ordem $\rho (\leq 0)$ regula a taxa de convergência em qualquer uma das condições de primeira ordem, em (3.21), e pode ser definido (ver Secção 2.2.4), como o parâmetro que surge na relação limite

$$\lim_{t \rightarrow \infty} \frac{\ln U(tx) - \ln U(t) - \xi \ln x}{A(t)} = \begin{cases} (x^{\rho} - 1)/\rho, & \text{se } \rho < 0, \\ \ln x, & \text{se } \rho = 0, \end{cases} \quad (3.22)$$

que é assumida para $x > 0$, e onde $|A|$ deve ser de variação regular com índice ρ . Esta condição tem sido amplamente aceite como uma condição apropriada para especificar a cauda direita de uma distribuição de tipo Pareto de forma semi-paramétrica. Por simplicidade admite-se que se lida com a classe de modelos Hall-Welsh (Hall and Welsh, 1985), com uma função quantil recíproca definida em (2.41).

Adaptações triviais dos resultados em de Haan and Peng (1998), Dekkers *et al.* (1989), Drees *et al.* (2004), Beirlant *et al.* (2005), Caeiro *et al.* (2005), de Haan and

Ferreira (2006), Fraga Alves *et al.* (2009) e Caeiro and Gomes (2011), respetivamente para os estimadores H, M, ML, GH, CH, GPPWM, MM e PPWM, permitem estabelecer o seguinte teorema.

Teorema 3.2.1. *Considere-se a validade da condição de segunda ordem em (3.22). Seja $k = k(n)$ uma sucessão intermédia de inteiros, i.e., verificando (2.13) e suponha-se que se trabalha com valores de k tais que $\lambda_A := \lim_{n \rightarrow \infty} \sqrt{k} A(n/k)$ é um valor finito. Então garante-se que os estimadores-EVI H, M, GH, MM, ML, para $\xi > 0$, e os estimadores-EVI PPWM, GPPWM, para $0 < \xi < 1/2$, genericamente denotados por $\hat{\xi}^\bullet(k)$, e respetivamente definidos em (3.2), (3.10), (3.11), (3.12), (3.20), (3.15) e (3.14), têm distribuição assintótica normal, i.e.*

$$\sqrt{k} \left(\hat{\xi}^\bullet(k) - \xi \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(\lambda_A b_\bullet, \sigma_\bullet^2), \quad (3.23)$$

onde $\mathcal{N}(\mu, \sigma^2)$ significa a distribuição normal com valor médio μ e variância σ^2 , e

$$b_H = \frac{1}{1-\rho}, \quad \sigma_H^2 = \xi^2, \quad b_M = b_{GH} = \frac{\xi - \xi\rho + \rho}{\xi(1-\rho)^2}, \quad \sigma_M^2 = \sigma_{GH}^2 = 1 + \xi^2,$$

$$b_{MM} = b_{ML} = \frac{(1+\xi)(\xi+\rho)}{\xi(1-\rho)(1+\xi-\rho)}, \quad \sigma_{MM}^2 = \sigma_{ML}^2 = (1+\xi)^2,$$

$$b_{PPWM} = \frac{(1-\xi)(2-\xi)}{(1-\xi-\rho)(2-\xi-\rho)}, \quad \sigma_{PPWM}^2 = \frac{\xi^2(1-\xi)(2-\xi)^2}{(1-2\xi)(3-2\xi)},$$

$$b_{GPPWM} = \frac{(\xi+\rho) b_{PPWM}}{\xi} \quad \text{e} \quad \sigma_{GPPWM}^2 = \frac{(1-\xi+2\xi^2)(1-\xi)(2-\xi)^2}{(1-2\xi)(3-2\xi)}.$$

Se ainda se admitir que se trabalha com a classe de modelos Hall-Welsh em (2.41), e se $(\hat{\beta}, \hat{\rho})$ for um estimador consistente de (β, ρ) , com $\hat{\rho} - \rho = o_{\mathbb{P}}(1/\ln n)$, com $\hat{\rho} - \rho = o_{\mathbb{P}}(1/\ln n)$, tem-se que $b_{CH} = 0$ e $\sigma_{CH}^2 = \sigma_H^2 = \xi^2$, para o estimador MVRB em (3.8).

Para os estimadores do EVI dependentes do parâmetro p , também, com as devidas alterações dos resultados em Gomes and Martins (2001), Caeiro and Gomes (2002a) e Brilhante *et al.* (2013), respetivamente para as classes de estimadores GM_p , $CG_{p,\delta}$ e H_p , tem-se o seguinte teorema.

Teorema 3.2.2. *Sob a validade da condição de primeira ordem, em (3.21), e supondo $k = k(n)$ uma sucessão intermédia de inteiros, i.e., verificando (2.13), as classes de estimadores GM_p , $CG_{p,\delta}$ e H_p respetivamente definidos em (3.16), (3.17) e (3.18) são*

consistentes para a estimação de $\xi > 0$, desde que $p \in \mathcal{R}_\bullet$, onde $\mathcal{R}_{H_p} = \{p : p < 1/\xi\}$, $\mathcal{R}_{GM_p} = \{p : p > -1\}$ e $\mathcal{R}_{CG_{p,\delta}} = \{(p, \delta) : p > 0, \delta > -1/p\}$.

Sob a validade das condições do Teorema 3.2.1, para qualquer dos estimadores (3.16), (3.17) e (3.18), também genericamente denotados por $\hat{\xi}^\bullet(k)$, e para regiões adequadas dos parâmetros p e δ , a equação (3.30) é válida, com

$$b_{GM_p} = \frac{1 - (1 - \rho)^p}{\rho(1 - \rho)^p} - \frac{p - 1}{1 - \rho}, \quad \sigma_{GM_p}^2 = \xi^2 \left\{ \frac{\Gamma(2p + 1)}{\Gamma^2(p + 1)} - p^2 \right\} \quad (p > -1/2), \quad (3.24)$$

$$b_{CG_{p,\delta}} = \frac{(1 - \rho)^{-\delta p} - \delta(1 - \rho)^{-p+1} + \delta - 1}{\delta \rho},$$

$$\sigma_{CG_{p,\delta}}^2 = \frac{\xi^2}{\delta^2} \left\{ \frac{2\Gamma(2\delta p)}{\delta p \Gamma^2(\delta p)} + \frac{\delta^2 \Gamma(2p - 1)}{\Gamma^2(p)} - \frac{2\Gamma((\delta + 1)p)}{p \Gamma(p) \Gamma(\delta p)} - (\delta - 1)^2 \right\} \quad (p > 1/2, \delta > 0), \quad (3.25)$$

$$b_{H_p} = \frac{1 - p\xi}{1 - p\xi - \rho} \quad \sigma_{H_p}^2 = \frac{\xi^2(1 - p\xi)^2}{1 - 2p\xi} \quad (p < 1/(2\xi)). \quad (3.26)$$

Além disso, se se trabalhar com a classe de modelos Hall-Welsh em (2.41), e se $(\hat{\beta}, \hat{\rho})$ for um estimador consistente de (β, ρ) , com $\hat{\rho} - \rho = o_p(1/\ln n)$, consegue-se que a componente dominante do viés de L_p^{RB} em (3.9) seja nula, i.e., $b_{L_p^{RB}} = 0$ e que a sua variância se mantenha igual à do correspondente estimador-EVI de viés não reduzido.

Mais especificamente do que o estabelecido nos dois teoremas anteriores, note-se que se se admitir que $k = k(n)$ é uma sucessão intermédia de inteiros e que se trabalha com valores de k tais que $\lambda_A := \lim_{n \rightarrow \infty} \sqrt{k} A(n/k)$ é um valor finito, a validade da condição de segunda ordem em (3.22), para regiões adequadas dos parâmetros p e δ permite escrever para todo o $\rho \leq 0$ a seguinte representação distribucional assintótica,

$$\hat{\xi}^\bullet(k) \stackrel{d}{=} \xi + \frac{\sigma_\bullet}{\sqrt{k}} Z_k^\bullet + b_\bullet A(n/k) + o_p(A(n/k)) \quad (3.27)$$

onde Z_k^\bullet é uma v.a. normal padrão e $(b_\bullet, \sigma_\bullet) \in \mathbb{R} \times \mathbb{R}^+$ são definidos nos dois teoremas anteriores. A $b_\bullet A(n/k)$ e σ_\bullet^2/k dá-se o nome respetivamente de *viés assintótico* e *variância assintótica* de $\hat{\xi}^\bullet(k)$.

Relativamente ao caso particular $\delta = 1$, em (3.17), i.e. para a classe de estimadores-EVI L_p , em (3.7), pode-se estabelecer o seguinte:

Corolário 3.2.1. *Sob a validade da condição de primeira ordem em (3.21), a classe de estimadores L_p , em (3.7), é consistente para a estimação de $\xi > 0$ desde que $p \in \mathcal{R}_L = \{p : p > 0\}$. Sob a validade das condições do Teorema 3.2.1, a equação (3.30) é válida, com*

$$b_{L_p} = \frac{1}{(1-\rho)^p} \quad \text{e} \quad \sigma_{L_p}^2 = \frac{\xi^2 \Gamma(2p-1)}{\Gamma^2(p)} \quad (p > 1/2) \quad (3.28)$$

e tem-se para $p > 1/2$ a seguinte representação distribucional,

$$L_p(k) \stackrel{d}{=} \xi + \frac{\sigma_{L_p} Z_k^{(p)}}{\sqrt{k}} + b_{L_p} A(n/k) + o_{\mathbb{P}}(A(n/k)) + o_{\mathbb{P}}(1/\sqrt{k}) \quad (3.29)$$

com $(b_{L_p}, \sigma_{L_p}^2)$ dado em (6.2), e onde $Z_k^{(p)}$ é uma v.a. com distribuição assintótica normal padrão.

Demonstração. Como já foi referido, a lei dos grandes números permite afirmar que as estatísticas em (3.7) são consistentes para a estimação de $\xi > 0$ para todo o $p > 0$. Com Y denotando uma vez mais uma v.a. com distribuição de Pareto standard e trabalhando sob a estrutura de segunda ordem em (3.22) com $\rho < 0$, tem-se,

$$\begin{aligned} \ln U(tx) - \ln U(t) - \xi \ln x - A(t) \frac{x^\rho - 1}{\rho} &= o(A(t)), \quad t \rightarrow \infty \iff \\ \ln \frac{U(tx)}{U(t)} &= \xi \ln x + A(t) \frac{x^\rho - 1}{\rho} + o(A(t)). \end{aligned}$$

Atendendo a que $X_{i:n} \stackrel{d}{=} U(Y_{i:n})$, $\frac{Y_{n-i+1:n}}{Y_{n-k:n}} \stackrel{d}{=} Y_{k-i+1:k}$ e $\frac{k}{n} Y_{n-k:n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 1$, e fazendo $t = \frac{n}{k} \rightarrow \infty$ (uma vez que $k = k(n)$, é uma sucessão intermédia de inteiros) e $x = Y_{k-i+1:k}$, tem-se,

$$\begin{aligned}
\frac{1}{k} \sum_{i=1}^k \left(\ln \frac{X_{n-i+1:n}}{X_{n-k:n}} \right)^p &\stackrel{d}{=} \frac{1}{k} \sum_{i=1}^k \left(\ln \frac{U(Y_{n-i+1:n})}{U(Y_{n-k:n})} \right)^p \\
&\stackrel{d}{=} \frac{1}{k} \sum_{i=1}^k \left(\ln \frac{U(Y_{n-k:n} Y_{k-i+1:k})}{U(Y_{n-k:n})} \right)^p \\
&\stackrel{d}{=} \frac{1}{k} \sum_{i=1}^k \left(\xi \ln Y_i + A(n/k) (Y_i^\rho - 1)/\rho + o_{\mathbb{P}}(A(n/k)) \right)^p,
\end{aligned}$$

quando $\frac{n}{k} \rightarrow \infty$.

Utilizando a aproximação,

$$(1+z)^p = 1 + pz + o(z), \text{ quando } z \rightarrow 0,$$

$$\frac{1}{k} \sum_{i=1}^k \left(\ln \frac{X_{n-i+1:n}}{X_{n-k:n}} \right)^p \stackrel{d}{=} \xi^p \frac{1}{k} \sum_{i=1}^k (\ln Y_i)^p + \frac{p \xi^{p-1} A(n/k)}{\rho} \frac{1}{k} \sum_{i=1}^k (\ln Y_i)^{p-1} (Y_i^\rho - 1) (1 + o_{\mathbb{P}}(1)).$$

Uma vez que $\ln Y \stackrel{d}{=} E$ e $Y \stackrel{d}{=} e^E$, com E uma v.a. com distribuição Exponencial padrão, a lei fraca dos grandes números permite escrever,

$$\frac{1}{k} \sum_{i=1}^k (\ln Y_i)^{p-1} (Y_i^\rho - 1) \xrightarrow[k \rightarrow \infty]{\mathbb{P}} \mathbb{E} [E^{p-1} (Y^\rho - 1)] = \Gamma(p) [(1-\rho)^{-p} - 1].$$

Consequentemente,

$$M_{k,n}^{(p)} \stackrel{d}{=} \xi^p \frac{1}{k} \sum_{i=1}^k E_i^p + \frac{\xi^{p-1} \Gamma(p+1) [(1-\rho)^{-p} - 1] A(n/k)}{\rho} + o_{\mathbb{P}}(A(n/k)).$$

Com a notação, $P_{k,n}^{(p)} := \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^k E_i^p - \Gamma(p+1) \right)$, pode-se escrever,

$$M_{k,n}^{(p)} \stackrel{d}{=} \xi^p \Gamma(p+1) \left\{ 1 + \frac{P_{k,n}^{(p)}}{\Gamma(p+1) \sqrt{k}} + \frac{(1-\rho)^{-p} - 1}{\xi \rho} A(n/k) + o_{\mathbb{P}}(A(n/k)) \right\}.$$

Então,

$$L_p(k) := \frac{M_{k,n}^{(p)}}{pM_{k,n}^{(p-1)}} \stackrel{d}{=} \xi \frac{\left\{ 1 + \frac{P_{k,n}^{(p)}}{\Gamma(p+1)\sqrt{k}} + \frac{(1-\rho)^{-p-1}}{\xi\rho} A(n/k) + o_{\mathbb{P}}(A(n/k)) \right\}}{\left\{ 1 + \frac{P_{k,n}^{(p-1)}}{\Gamma(p)\sqrt{k}} + \frac{(1-\rho)^{-(p-1)-1}}{\xi\rho} A(n/k) + o_{\mathbb{P}}(A(n/k)) \right\}}.$$

Tendo em conta as seguintes aproximações,

$$\frac{1}{1+x} = 1 - x + o(x), \text{ quando } x \rightarrow 0;$$

$$(1+x)(1+y) = 1 + x + y + o(xy), \text{ quando } x, y \rightarrow 0,$$

tem-se para $p > 1/2$,

$$L_p(k) \stackrel{d}{=} \xi \left\{ 1 + \frac{P_{k,n}^{(p)}}{\Gamma(p+1)\sqrt{k}} - \frac{P_{k,n}^{(p-1)}}{\Gamma(p)\sqrt{k}} + \frac{(1-\rho)^{-p}}{\xi} A(n/k) + o_{\mathbb{P}}(1/\sqrt{k}) + o_{\mathbb{P}}(A(n/k)) \right\}.$$

Note-se que,

$$\mathbb{E} \left[P_{k,n}^{(p)} \right] = 0 \text{ e } \mathbb{V}\text{ar} \left[P_{k,n}^{(p)} \right] = \Gamma(2p+1) - \Gamma^2(p+1),$$

$$\begin{aligned} \mathbb{C}\text{ov}(P_{k,n}^{(p-1)}, P_{k,n}^{(p)}) &= \mathbb{C}\text{ov} \left(\sqrt{k} \left(\frac{1}{k} \sum_{i=1}^k E_i^{p-1} - \Gamma(p) \right), \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^k E_i^p - \Gamma(p+1) \right) \right) \\ &= k \mathbb{C}\text{ov} \left(\frac{1}{k} \sum_{i=1}^k E_i^{p-1}, \frac{1}{k} \sum_{i=1}^k E_i^p \right) = \frac{\sum_{i=1}^k \mathbb{C}\text{ov}(E_i^{p-1}, E_i^p)}{k} \\ &= \mathbb{E}(E^{2p-1}) - \mathbb{E}(E^{p-1}) \mathbb{E}(E^p) = \Gamma(2p) - \Gamma(p)\Gamma(p+1), \end{aligned}$$

pelo que,

$$\begin{aligned} \mathbb{V}\text{ar} \left(\frac{P_{k,n}^{(p)}}{\Gamma(p+1)} - \frac{P_{k,n}^{(p-1)}}{\Gamma(p)} \right) &= \\ &= \frac{\Gamma(2p+1) - \Gamma^2(p+1)}{\Gamma^2(p+1)} + \frac{\Gamma(2p-1) - \Gamma^2(p)}{\Gamma^2(p)} - \frac{2(\Gamma(2p) - \Gamma(p)\Gamma(p+1))}{\Gamma(p)\Gamma(p+1)} = \frac{\Gamma(2p-1)}{\Gamma^2(p)} \end{aligned}$$

desde que $p > 1/2$.

Fazendo

$$Z_k^{(p)} := \frac{\frac{P_{k,n}^{(p)}}{\Gamma(p+1)} - \frac{P_{k,n}^{(p-1)}}{\Gamma(p)}}{\sqrt{\frac{\Gamma(2p-1)}{\Gamma^2(p)}}},$$

obtem-se a representação distribucional assintótica em (6.1), com (b_{L_p}, σ_{L_p}) dados em (6.2) e onde $Z_k^{(p)}$ é uma v.a. com distribuição assintótica normal padrão. Considerando que se trabalha com valores de k tais que $\lambda_A := \lim_{n \rightarrow \infty} \sqrt{k} A(n/k)$ é um valor finito, tem-se também,

$$\sqrt{k} \left(\hat{\xi}^{L_p}(k) - \xi \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N} \left(\lambda_A b_{L_p}, \sigma_{L_p}^2 \right). \quad (3.30)$$

□

Observação 3.2.1. *Um resultado similar para $p = 1$, i.e., para o estimador-EVI H , foi deduzido em de Haan and Peng (1998).*

Observação 3.2.2. *Note-se que, em relação aos estimadores-EVI L_p , em (3.7), o Corolário 3.2.1 é um caso particular do Teorema 1 em Caeiro and Gomes (2002a), mas incluindo generalizações triviais da consistência para $p > 0$, e normalidade assintótica para $p > 1/2$, em vez de $p \geq 1$. Note-se ainda que existe um total acordo entre este resultado e o fornecido no Teorema 1 de Caeiro and Gomes (2002a).*

Observação 3.2.3. *Note-se que uma alternativa ao estimador-EVI H , e relacionado com o estimador-EVI em (3.7), mas com $p = 2$, foi considerada em de Haan and Peng (1998), e aí atribuída a Casper de Vries. Para este estimador, $L_2(k) = M_{k,n}^2 / (2M_{k,n}^{(1)})$, Gomes et al. (2000) deduziu a representação distribucional assintótica*

$$L_2(k) \stackrel{d}{=} \xi + \frac{\xi \sqrt{2} Z_k^{(2)}}{\sqrt{k}} + \frac{A(n/k)}{(1-\rho)^2} + o_{\mathbb{P}}(A(n/k)) + o_{\mathbb{P}}(1/\sqrt{k}),$$

onde $Z_k^{(2)}$ tem distribuição assintótica normal padrão. Nesse artigo é feita também uma comparação assintótica no nível ótimo entre L_2 e L_1 .

Observação 3.2.4. *Em Gomes and Martins (2001) foi introduzida a classe de estimadores-EVI em (3.16), tendo sido realizado para esses estimadores um estudo ao nível assintótico e um também para amostras finitas. Para este estimador, Gomes and Martins (2001) obtiveram uma representação assintótica em distribuição do mesmo tipo da*

conseguida em (6.1) para $p > 0$, com $(\sigma_{L_p}^2, b_{L_p})$ substituído por $(\sigma_{GM_p}^2, b_{GM_p})$, em (3.24). De referir ainda que para esta classe de estimadores-EVI, existe um valor p_0 não explícito para o qual $b_{GM_{p_0}} = 0$, i.e, esta classe de estimadores pode ser de viés reduzido de segunda ordem.

Observação 3.2.5. Note-se ainda que, para a classe de estimadores-EVI MO_p , aqui denotados por H_p e definidos em (3.18), também se verifica, para $p < 1/(2\xi)$, uma representação distribucional do tipo da que se obtém em (6.1), com $(\sigma_{H_p}^2, b_{H_p})$ dados em (3.26).

Relativamente à classe de estimadores-EVI L_p , tem-se que, para qualquer $\xi > 0$, $\sigma_{L_p}^2(\xi)$, em (6.2), tem um mínimo em $p = 1$. Na Figura 3.1(a), exibe-se $\sigma_{L_p}(\xi)/\xi$, independente de ξ , como uma função de p . Na Figura 3.1(b), mostra-se que o comportamento de $b_{L_p}(\rho)$, em (6.2), independente de ξ , é para qualquer valor de ρ apresentado, sempre decrescente com p . Observa-se também que $b_{L_p}(\rho)$ apresenta valores mais baixos para valores de ρ em valor absoluto mais altos.

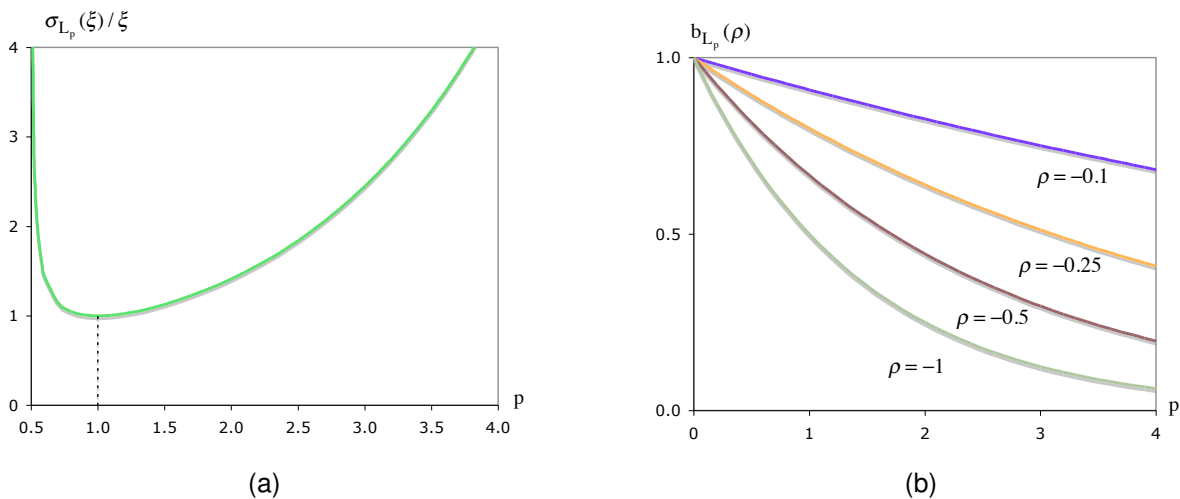


Figura 3.1: Gráfico de $\sigma_{L_p}(\xi)/\xi$, como função de $p > 0.5$ (a) e comportamento do viés assintótico $b_{L_p}(\rho)$, para $\rho = -0.1, -0.25, -0.5$ e -1 , como função de $p \geq 0$ (b).

Os resultados acima mencionados requerem uma comparação assintótica no nível k ou (k, p) , da classe de estimadores-EVI em estudo, um tópico a ser estudado na Secção 3.3.

3.3 Comparação assintótica no nível ótimo

Nesta secção procede-se à comparação assintótica dos estimadores-EVI nos seus níveis ótimos. Serão considerados todos os estimadores-EVI não dependentes de p e os estimadores-EVI H_p , (3.18), e L_p , (3.7), determinados no nível ótimo p . As classes GM_p , $CG_{p,\delta}$ $\delta > 1$ e L_p^{RB} respetivamente em (3.16), (3.17) e em (3.9) serão excluídas desta análise devido ao facto destas classes de estimadores-EVI poderem ser SORB em todo o plano- (ξ, ρ) .

Este estudo será realizado de forma similar ao que foi feito em de Haan and Peng (1998), Gomes and Martins (2001), Gomes *et al.* (2005, 2007b, 2013b, 2015b), Gomes and Neves (2008), Gomes and Henriques-Rodrigues (2010) e Brilhante *et al.* (2013).

Suponha-se que $\hat{\xi}^\bullet(k)$ denota qualquer um dos estimadores-EVI semi-paramétricos, para os quais se tem a representação distribucional assintótica,

$$\hat{\xi}^\bullet(k) = \xi + \frac{\sigma_\bullet Z_k^\bullet}{\sqrt{k}} + b_\bullet A(n/k) + o_{\mathbb{P}}(A(n/k)), \quad (3.31)$$

para qualquer sucessão intermédia de inteiros $k = k_n$, e onde Z_k^\bullet é uma v.a. assintoticamente normal padrão. Então, $\sqrt{k}(\hat{\xi}^\bullet(k) - \xi) \xrightarrow{d} N(\lambda_A b_\bullet, \sigma_\bullet^2)$ desde que k seja tal que $\sqrt{k} A(n/k) \rightarrow \lambda_A$, finito, com $n \rightarrow \infty$, podendo-se escrever $\text{Viés}_\infty(\hat{\xi}^\bullet(k)) := b_\bullet A(n/k)$ (viés assintótico) e $\text{Var}_\infty(\hat{\xi}^\bullet(k)) := \sigma_\bullet^2/k$ (variância assintótica).

Portanto, o chamado *erro quadrático médio assintótico* (AMSE, do inglês "asymptotic mean square error") é dado por,

$$\text{AMSE}(\hat{\xi}^\bullet(k)) := \sigma_\bullet^2/k + b_\bullet^2 A^2(n/k).$$

A teoria de variação regular (Bingham *et al.*, 1989) permitiu a Dekkers and de Haan (1993) mostrar que, sempre que $b_\bullet \neq 0$, existe uma função $\varphi(n) = \varphi(n, \xi, \rho)$, tal que,

$$\lim_{n \rightarrow \infty} \varphi(n) \text{AMSE}(\hat{\xi}_{n0}^\bullet) = (\sigma_\bullet^2)^{-\frac{2\rho}{1-2\rho}} (b_\bullet^2)^{\frac{1}{1-2\rho}} =: \text{LMSE}(\hat{\xi}_{n0}^\bullet),$$

onde $\hat{\xi}_{n0}^\bullet := \hat{\xi}^\bullet(k_{0|\bullet}(n))$ e $k_{0|\bullet}(n) := \arg \min_k \text{MSE}(\hat{\xi}^\bullet(k))$. Além disso, se se restringir a condição de segunda ordem (3.22) assumindo que $A(t) = \xi \beta t^\rho$, $\rho < 0$, tal como

sucede para a classe referida em (2.41), pode-se escrever,

$$k_{0|\bullet}(n) := \arg \min_k \text{AMSE} \left(\hat{\xi}^\bullet(k) \right) = \left(\frac{\sigma_\bullet^2 n^{-2\rho}}{b_\bullet^2 \xi^2 \beta^2(-2\rho)} \right)^{1/(1-2\rho)} (1 + o(1)).$$

E também considerar a seguinte:

Definição 3.3.1. *Dados dois estimadores com viés, $\hat{\xi}^{(1)}(k)$ e $\hat{\xi}^{(2)}(k)$, para os quais a representação distribucional do tipo da especificada em (3.31) se verifica, com constantes (σ_1, b_1) e (σ_2, b_2) , $b_1, b_2 \neq 0$, respetivamente, ambos determinados nos correspondentes níveis ótimos, a raiz da eficiência assintótica (AREFF, do inglês "asymptotic root efficiency") de $\hat{\xi}_{n0}^{(1)}$ relativamente a $\hat{\xi}_{n0}^{(2)}$ é definida como,*

$$\text{AREFF}_{1|2} \equiv \text{AREFF}_{\hat{\xi}_{n0}^{(1)}|\hat{\xi}_{n0}^{(2)}} := \sqrt{\text{LMSE}(\hat{\xi}_{n0}^{(2)})/\text{LMSE}(\hat{\xi}_{n0}^{(1)})} = \left(\left(\frac{\sigma_2}{\sigma_1} \right)^{-2\rho} \left| \frac{b_2}{b_1} \right| \right)^{\frac{1}{1-2\rho}}. \quad (3.32)$$

Observação 3.3.1. *Note-se que o indicador AREFF, em (3.32), foi concebido de forma a que quanto maior for o valor do indicador AREFF melhor será a eficiência do primeiro estimador, denotado pelo sobrescrito ⁽¹⁾.*

Observação 3.3.2. *As classes GM_p , $\text{CG}_{p,\delta}$ $\delta > 1$, respetivamente em (3.16) (3.17), podem ser SORB em todo o plano- (ξ, ρ) , assim como acontece com a classe L_p^{RB} em (3.9). Por conseguinte serão excluídas da análise de comparação assintótica no nível ótimo, uma vez que não se pode aplicar a **Definição 3.3.1**.*

Observação 3.3.3. *Note-se ainda que os estimadores-EVI MM, GPPWM e ML, em (3.12), (3.14) e (3.20), respetivamente, são todos estimadores SORB na região $\xi + \rho = 0$ (onde $b_{\text{MM}} = b_{\text{GPPWM}} = b_{\text{ML}} = 0$). Portanto, também para estes estimadores e nessas regiões, não se pode aplicar a **Definição 3.3.1**. Relativamente a estes estimadores, pelo facto de poderem ser SORB, espera-se que assintoticamente estes estimadores superem, no nível ótimo, qualquer outro estimador-EVI que não seja SORB. No entanto, $\sigma_{\text{GPPWM}} > \sigma_{\text{MM}} = \sigma_{\text{ML}} > \sigma_{\text{H}} = \sigma_{\text{CH}}$ e por isso tudo depende da componente dominante do viés e da variância assintótica, ou seja tudo depende do estudo comparativo do comportamento dos erros quadráticos médios, que é sem dúvida um desafio*

para investigação futura, fora do âmbito deste trabalho, parcialmente tratado em Ca-
eiro *et al.* (2009, 2016a). Um comentário semelhante aplica-se ao comportamento dos
estimadores-EVI M e ao GH na região $\xi = -\rho/(1-\rho)$ (onde $b_M = b_{GH} = 0$), mas tendo
uma variância assintótica igual a $1 + \xi^2 > \xi^2$, a variância assintótica de H e CH .

Observação 3.3.4. Os estimadores-EVI L_p e H_p , respetivamente dados em (3.7) e
(3.18), nunca são estimadores SORB, e por isso serão crucialmente incluídos no es-
tudo de comparação assintótica da Secção 3.3.1.

3.3.1 Comparação assintótica dos estimadores-EVI L_p e H_p nos níveis ótimos

Considere-se a classe de estimadores-EVI $L_p(k)$ em (3.7). Tem-se,

$$\text{LMSE}(L_{0|p}) = (\xi^2 \Gamma(2p-1)/\Gamma^2(p))^{-\frac{2\rho}{1-2\rho}} ((1-\rho)^{-2p})^{\frac{1}{1-2\rho}}$$

e

$$\text{AREFF}_{L_p|L_1} = \left(\left(\Gamma(p)/\sqrt{\Gamma(2p-1)} \right)^{-2\rho} (1-\rho)^{p-1} \right)^{\frac{1}{1-2\rho}}. \quad (3.33)$$

Observação 3.3.5. Em Gomes *et al.* (2000) mostrou-se que a a raiz da eficiência rela-
tiva assintótica (AREFF) de $L_2(k)$ comparativamente a $L_1(k)$ é dada por $\text{AREFF}_{L_2|L_1} =$
 $[2^\rho(1-\rho)]^{1/(1-2\rho)}$, de acordo com (3.33). Conforme observado no referido artigo,
 $\text{AREFF}_{L_2|L_1} > 1 \iff -1 < \rho < 0$.

Para medir a performance de $H_{0|p}$, com H_p o estimador-EVI $MO_p(k)$ em (3.18),
Brilhante *et al.* (2013) calcularam um indicador- AREFF semelhante, dado por,

$$\text{AREFF}_{H_p|H_0} = \left(\left(\frac{\sqrt{1-2p\xi}}{1-p\xi} \right)^{-2\rho} \left| \frac{1-p\xi-\rho}{(1-\rho)(1-p\xi)} \right| \right)^{\frac{1}{1-2\rho}}, \quad (3.34)$$

reparametrizado em $(\rho, a = p\xi < 1/2)$, e denotado por $\text{AREFF}_{a|0}^*$. A Figura 3.2
corresponde aos gráficos de contorno de $\text{AREFF}_{L_p|L_1}$ em (3.33) (3.2(a)) e de $\text{AREFF}_{a|0}^*$
em (3.34) (3.2(b)).

O ganho em eficiência não é muito elevado, mas, nos níveis ótimos, existe uma
alargada região do plano- (p, ρ) onde a classe de estimadores-EVI L_p tem um desem-
penho melhor do que o estimador de Hill, com uma eficiência ligeiramente superior

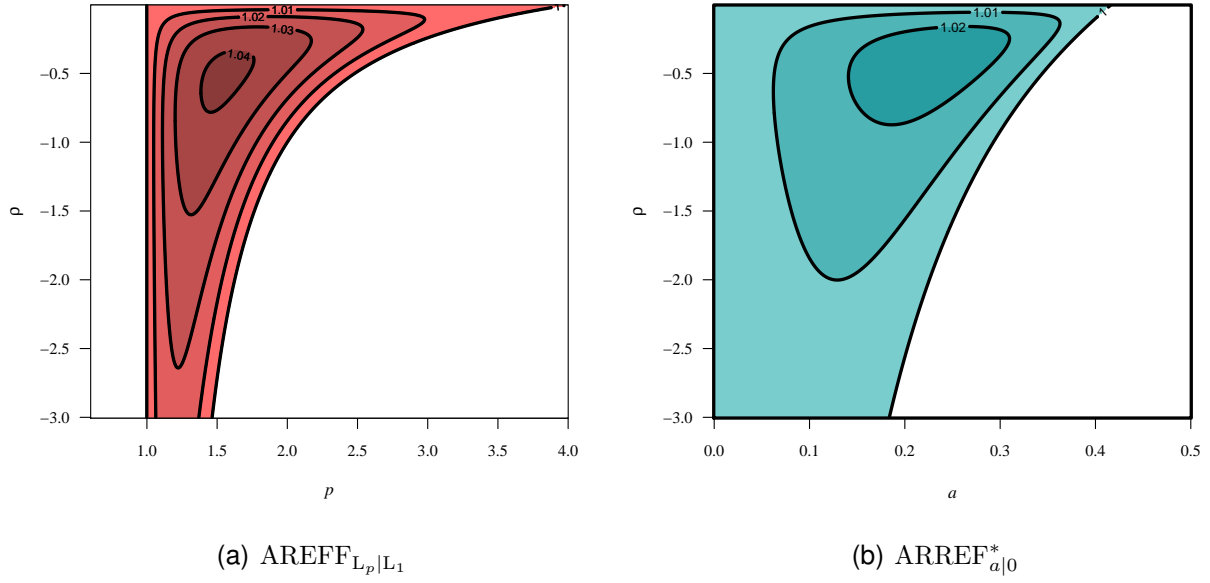


Figura 3.2: Gráfico de contornos.

àquela associada com a comparação de H_p e o Hill, no plano- (a, ρ) . Este resultado em conjunto com o facto que, tanto quanto se conhece, o estimador-EVI em (3.18) determinado no nível ótimo p , no sentido de maximização de $\text{AREFF}_{H_p|H_0}$, i.e., em $p_{M|H} \equiv p_{M|H}(\rho) := \arg \max_p \text{AREFF}_{H_p|H_0}$, explicitamente dado por,

$$p_{M|H} = \varphi_\rho / \xi, \quad \text{com} \quad \varphi_\rho := 1 - \rho/2 - \sqrt{\rho^2 - 4\rho + 2} / 2$$

com $b_{p_{M|H}} \neq 0$, é o único estimador-EVI que não sendo RB consegue ter uma eficiência melhor do que o estimador-EVI Hill em todo o plano- (ξ, ρ) , imediatamente se coloca a questão sobre o que acontece para o valor ótimo de p associado com o estimador-EVI L_p .

Na Figura 3.3(a) representa-se o indicador $\text{AREFF}_{L_p|L_1}$, como função de p para $\rho = -3(0.1)0$ e o valor de $p_{M|L} = p_{M|L}(\rho) := \arg \max_p \text{AREFF}_{L_p|L_1}$, como função de ρ (Figura 3.3(b)).

Na verdade, assim como $\text{AREFF}_{H_{p_{M|H}}|H_0} > 1$, para qualquer $\rho < 0$ e $\xi > 0$, também, em $p_{M|L} \equiv p_{M|L}(\rho) := \arg \max_p \text{AREFF}_{L_p|L_1}$,

$$\text{AREFF}_{L_{p_{M|L}}|L_1} > 1,$$

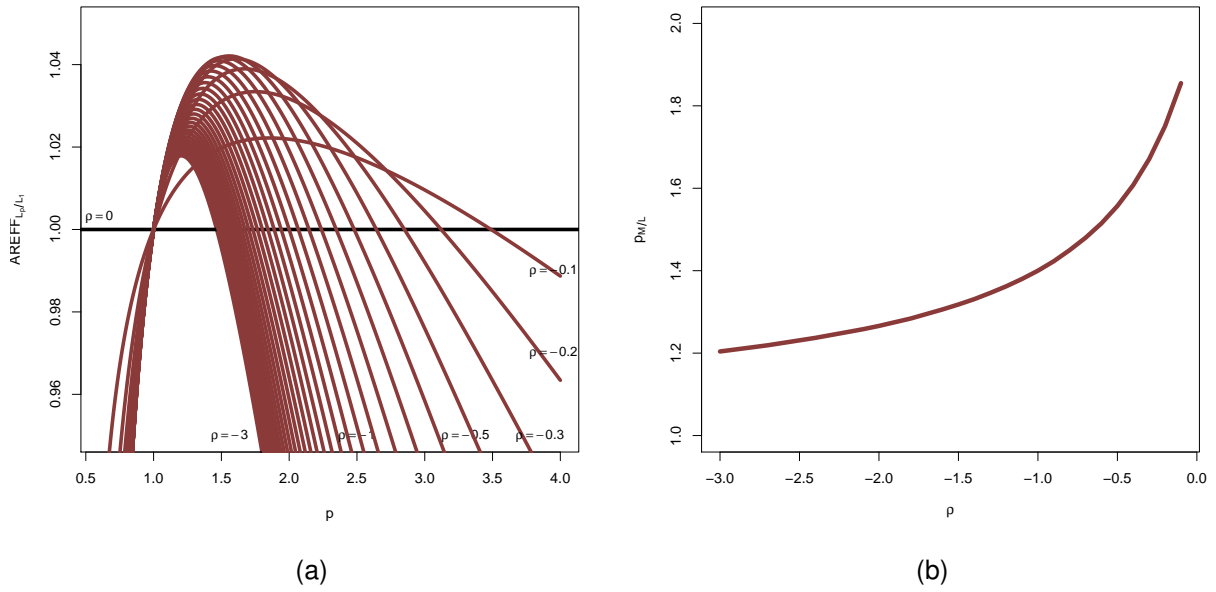


Figura 3.3: $\text{AREFF}_{L_p|L_1}$, como função de p , para $\rho = -3(0.1)0$ (a) e o valor de $p_{M|L} = p_{M|L}(\rho)$, como função ρ (b)

para qualquer $\rho < 0$ e $\xi > 0$. Aliás,

$$\text{AREFF}_{L_{p_{M|L}}|L_1} > \text{AREFF}_{H_{p_{M|H}}|H_0},$$

tal como é ilustrado na Figura 3.4.

3.3.2 Comparação global dos estimadores-EVI nos níveis ótimos

A Figura 3.5 apresenta uma comparação da eficiência, no nível ótimo, de todos os "clássicos" estimadores-EVI em consideração, incluindo ambas as classes L_p e H_p , realizada sob o critério do indicador (AREFF). Note-se que esta comparação apenas contempla os estimadores-EVI que não são RB (ou em regiões para as quais $b_{\bullet} \neq 0$). Assim a Figura 3.5 apresenta unicamente os estimadores que são mais eficientes em regiões do plano (ξ, ρ) nas quais os estimadores não verifiquem $b_{\bullet} = 0$. Como mencionado atrás e detetado pela primeira vez por Brilhante *et al.* (2013), no nível ótimo o estimador-EVI H_p pode bater em eficiência o estimador-EVI Hill em todo o plano- (ξ, ρ) . Mas é agora ultrapassado pelo estimador-EVI Lehmer, no nível ótimo, também em todo o plano- (ξ, ρ) , um comportamento atípico entre os estimadores-EVI clássicos. No

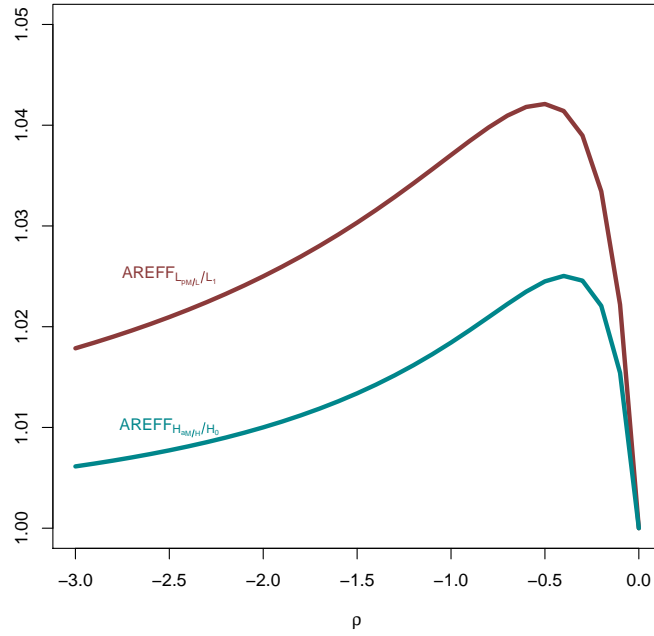


Figura 3.4: $\text{AREFF}_{L_{p_M|L}|L_1}$ e $\text{AREFF}_{H_{p_M|H}|H_0}$, como função de $\rho = -3(0.1)0$

entanto, o estimador-EVI L, no nível ótimo, pode ser ultrapassado em eficiência pelo estimador-EVI M, em (3.10), ou GH, em (3.11), numa região próxima a $\xi = -\rho/(1-\rho)$, onde $b_M = b_{GH} = 0$. O estimador-MM, em (3.12), assintoticamente equivalente ao estimador-ML, em (3.20), exceto quando $\xi + \rho = 0$ e $(\xi, \rho) \neq (0, 0)$, supera em eficiência, nos níveis ótimos, o estimador-M, numa região em redor de $\xi + \rho = 0$, e até superando o estimador-EVI L no nível ótimo. O estimador-EVI GPPWM é um estimador RB para $\xi + \rho = 0$ e pode ser mais eficiente do que o estimador-EVI MM numa pequena região do plano- (ξ, ρ) . De referir ainda que o estimador-EVI PPWM é mais eficiente para valores de ξ em torno de 0.05.

Como seria esperado, nenhum dos estimadores domina as alternativas, mas a classe de estimadores-EVI L_p tem um bom desempenho, superando, nos níveis ótimos, a classe de estimadores-EVI $MO_p \equiv H_p$ em todo o plano- (ξ, ρ) .

[illegible]

Figura 3.5: Indicação dos estimadores-EVI mais eficientes calculados nos níveis ótimos.

Capítulo 4

Propriedades das Classes de Estimadores de Lehmer em Amostras de Dimensão Finita

Neste capítulo realiza-se um estudo detalhado do comportamento das classes de estimadores-EVI de Lehmer, L_p em (3.7) e L_p^{RB} em (3.9), para amostras de dimensão finita, comparando com o comportamento dos outros estimadores-EVI em consideração.

Este estudo baseia-se na simulação multi-amostra Monte Carlo de dimensão $50000 = 5000(\text{"run"}) \times 10(\text{réplicas})$. Desta forma, para cada réplica são geradas 5000 amostras de dimensão n , com $n = 500$, e 1000, subjacentes aos seguintes modelos pertencentes à classe de Hall-Welsh, em (2.41):

1. o modelo *Fréchet*, com f.d. $F(x) = \exp(-x^{-1/\xi})$, $x \geq 0$, para valores de $\xi = 0.1, 0.5, 1, 2$ e 3 ($\rho = -1$);
2. o modelo de *valores extremos*, com f.d. $F(x) = EV_\xi(x)$, com $EV_\xi(x)$ dado em (2.17), para os mesmos valores de ξ ($\rho = -\xi$, $\xi < 1$; $\rho = -1$, $\xi \geq -1$). Note-se que neste caso apenas se consideram os valores positivos da amostra, pelo que as amostras foram geradas com uma dimensão superior às referidas.
3. o modelo *Burr*, com f.d. $F(x) = 1 - (1 + x^{-\rho/\xi})^{-1/\rho}$, $x > 0$, $\xi > 0$ e $\rho < 0$, para valores de $\xi = 0.1, 0.5, 1, 2$ e 3 e valores de $\rho = -0.1, -0.5, -2$ e -3 .

O presente estudo de simulação permite obter para os modelos referidos, e para os estimadores $\hat{\xi}^\bullet(k)$ em análise, as estimativas,

- do valor médio, $\widehat{E}[\hat{\xi}^\bullet(k)]$;
- da raiz do erro quadrático médio (RMSE, do inglês "*root mean square error*"), $\widehat{\text{RMSE}}[\hat{\xi}^\bullet(k)]$;
- $k_{0|\bullet} = \arg \min_k \widehat{\text{RMSE}}[\hat{\xi}^\bullet(k)]$.

Também foi simulada uma medida de eficiência de $\hat{\xi}^\bullet$ relativamente a $\hat{\xi}^H$, em (3.2), nos níveis ótimos, através do indicador,

$$\text{REFF}_{\hat{\xi}^\bullet|\hat{\xi}^H} := \frac{\widehat{\text{RMSE}}[\hat{\xi}^H(k_0)]}{\widehat{\text{RMSE}}[\hat{\xi}^\bullet(k_0)]}. \quad (4.1)$$

Note-se que quanto maior for o valor dessa medida melhor será a eficiência de $\hat{\xi}^\bullet$ relativamente a $\hat{\xi}^H$.

De referir também que neste estudo de simulação as 10 réplicas possibilitam o cálculo de intervalos de confiança.

4.1 Padrões do valor médio e do erro quadrático médio dos estimadores-EVI

Numa primeira fase, para $n = 1000$ e para cada um dos modelos acima mencionados, foi simulado o valor médio (E) e a raiz do erro quadrático médio (RMSE) dos estimadores $L_p(k)$, em (3.7), e L_p^{RB} em (3.9), como funções do número de e.o. de topo k envolvidas na estimação e para alguns valores de $p \geq 1$. Foi realizada também uma comparação destes estimadores com outros estimadores já referenciados ao longo do trabalho. Neste âmbito consideraram-se, para além das classes de estimadores de Lehmer em estudo, e do estimador de Hill em (3.2), todos os estimadores que se distinguiram no estudo comparativo assintótico nos níveis ótimos, o estimador $M(k)$ em (3.10) e o $MM(k)$ em (3.12). Consideraram-se também a classe de estimadores

$H_p(k)$ em (3.18) e os estimadores que podem ser de viés reduzido, em todo o plano- (ξ, ρ) , os estimadores $GM_p(k)$ em (3.16). Note-se que relativamente aos estimadores que dependem do parâmetro p , teve-se em conta os valores de p para os quais esses estimadores são definidos e para os quais se verifica o limite em distribuição para a normal. Assim, no caso da classe de estimadores GM_p , consideraram-se alguns valores de $p \geq 0$ e para a classe de estimadores H_p , valores de $p \in [0, 1/(2\xi)[$.

Os gráficos que se seguem representam as trajetórias das estimativas simuladas de E e de RMSE para cada um dos modelos referidos. Note-se que relativamente aos estimadores GM_p e H_p , apenas se apresentam essas estimativas para valores de p para os quais o desempenho dos estimadores parece ser melhor.

Para obtenção dos resultados de simulação foram desenvolvidos e implementados algoritmos no *software* R e construídas as funções necessárias.

Modelo Fréchet $(\xi)(\rho = -1)$

Relativamente a este modelo verifica-se que, de acordo com as trajetórias do RMSE, e independentemente do índice de valores extremos ξ , os valores de p , entre 2 e 2.5, para a classe de estimadores $L_p(k)$ (como ilustração ver Figuras 4.1, e no Anexo A, as Figuras A.1, A.2, A.3 e A.4), e entre 1 e 2, para a classe $L_p^{RB}(k)$ (ver Figuras 4.2, e no Anexo A, as Figuras A.5, A.6, A.7 e A.8), parecem melhorar o desempenho desses estimadores, mesmo numa vasta região de valores de k . Note-se que, para esses valores de p , essas trajetórias exibem uma forma de "banheira", resultando em valores baixos de RMSE numa vasta região de valores de k , o que é aconselhável. Como seria esperado, e como se pode ver nas correspondentes Figuras relativas a E , também se observa que os estimadores $L_p^{RB}(k)$ exibem uma estabilidade em torno do verdadeiro valor do parâmetro muito superior àquela que se observa nos estimadores $L_p(k)$, sendo o estimador $CH \equiv L_1^{RB}(k)$ aquele que apresenta maior estabilidade.

As Figuras 4.3, 4.4, 4.5, 4.6 e 4.7 mostram, para $\xi = 0.5, 2, 3$, uma performance melhor nos estimadores que são RB, $L_1^{RB}(k) \equiv CH$ e $L_2^{RB}(k)$, e no estimador $GM_{2.5}$, que nestas condições, $\rho = -1$ e $p = 2.5$, tem um viés assintótico, $b_{GM_p} = 0.07$, próximo

de zero. Para além destes estimadores, também o estimador MM, quando $\xi = 1$, se revela um estimador eficiente (note-se que quando $\xi = 1$ e $\rho = -1$ este estimador é RB). De referir ainda que, para $\xi = 0.1, 0.5$, os estimadores $M(k)$ e $MM(k)$ não parecem comportar-se bem, sendo sempre superados pelo estimador $L_{2.5}(k)$.

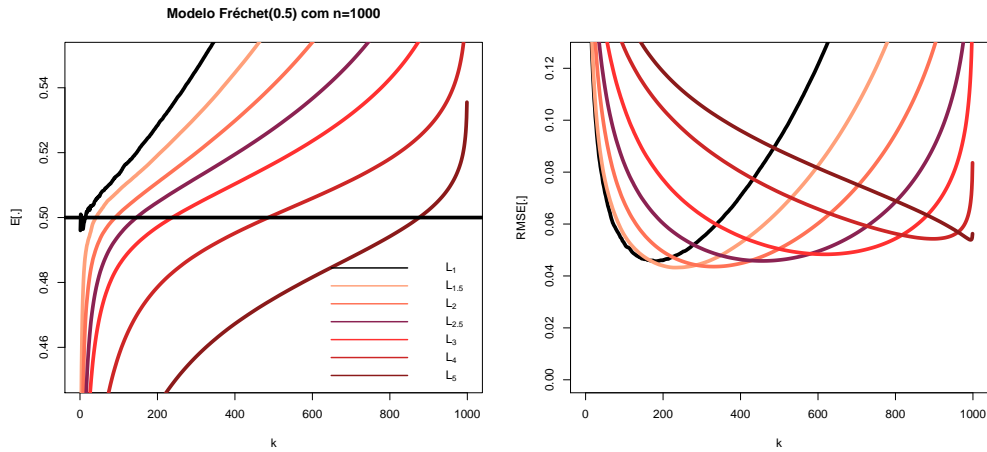


Figura 4.1: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1(0.5)3, 4, 5$.

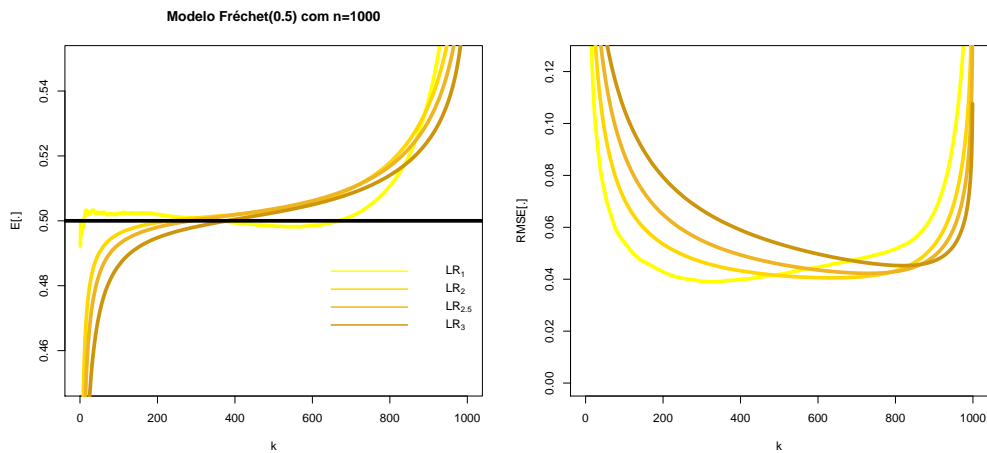


Figura 4.2: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{RB} , $p = 1, 2, 2.5, 3$.

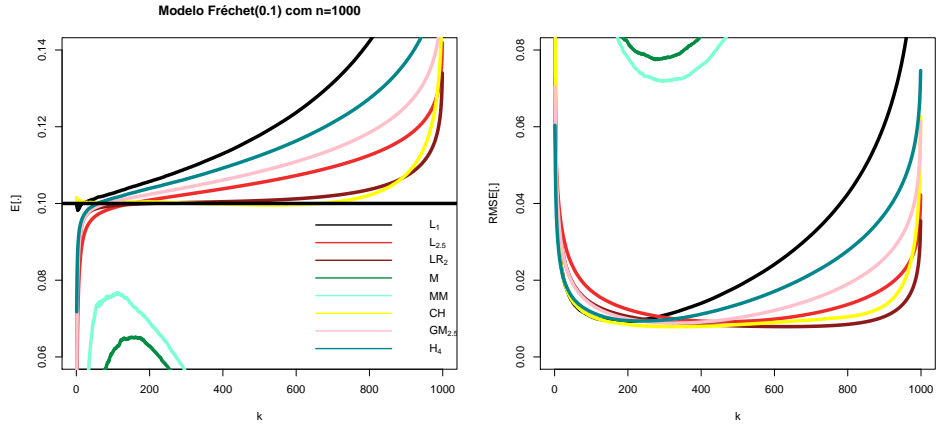


Figura 4.3: Comparação das estimativas simuladas de E e de RMSE.

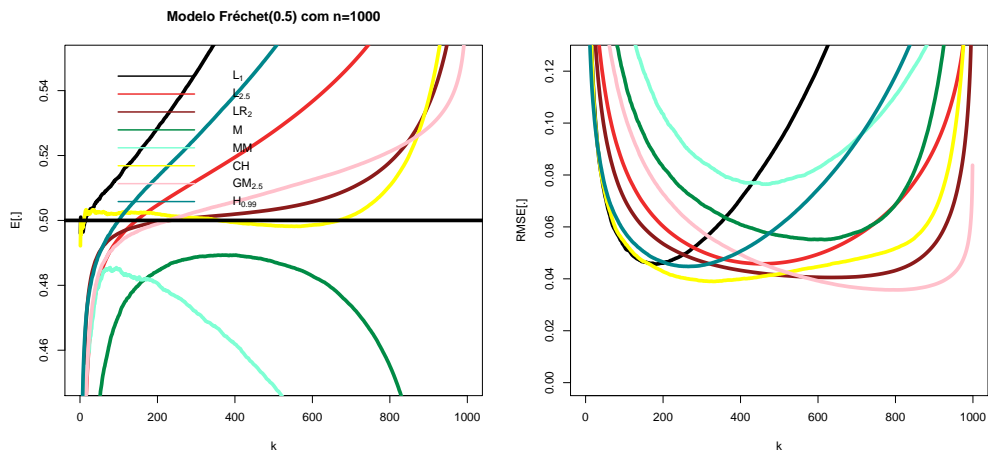


Figura 4.4: Comparação das estimativas simuladas de E e de RMSE.

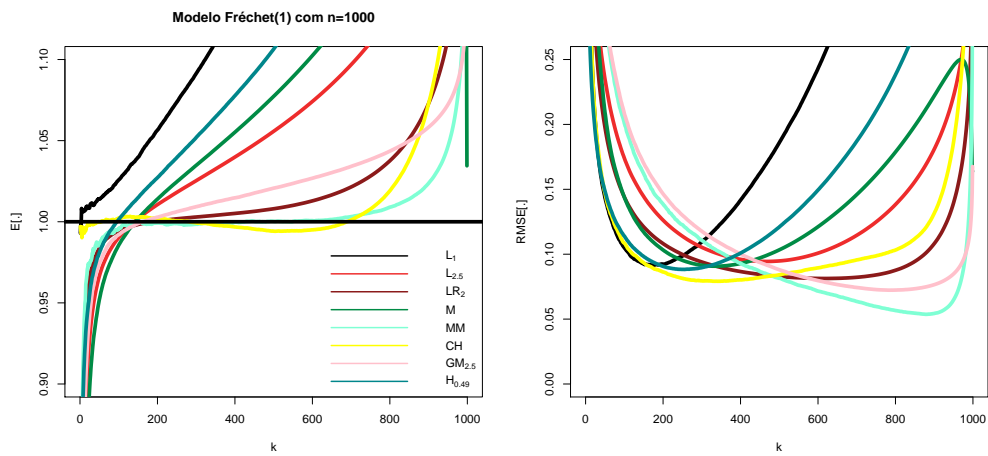


Figura 4.5: Comparação das estimativas simuladas de E e de RMSE.

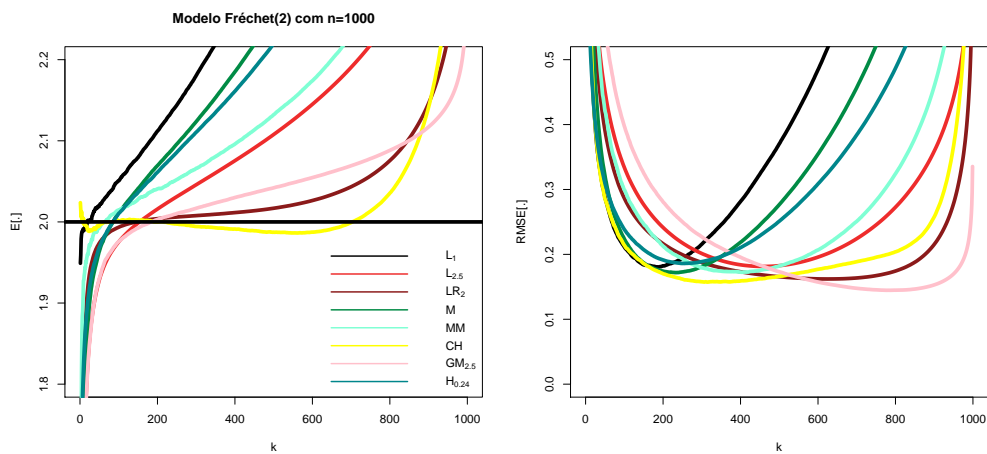


Figura 4.6: Comparação das estimativas simuladas de E e de RMSE.

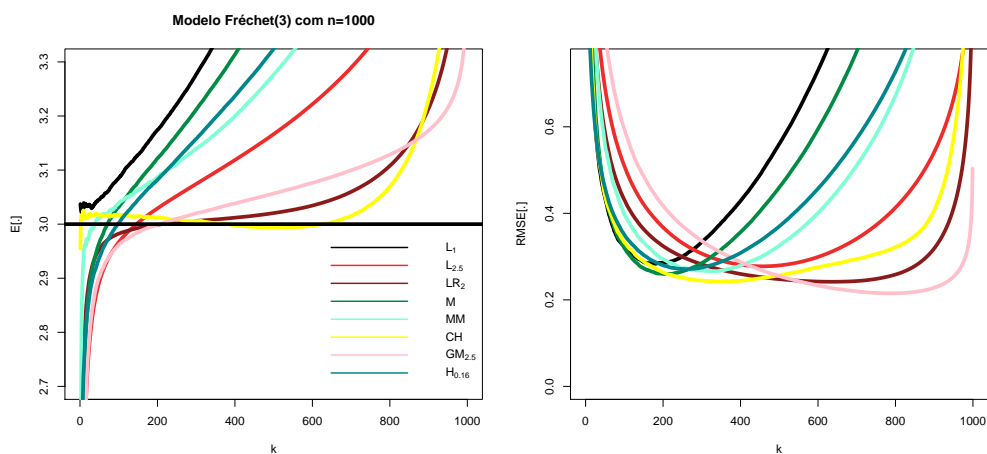


Figura 4.7: Comparação das estimativas simuladas de E e de RMSE.

Modelo $EV(\xi)(\xi = 1, 2, 3 \text{ e } \rho = -1)$

Independentemente do valor de ξ , os valores de p que tornam os estimadores $L_p(k)$ e $L_p^{RB}(k)$ mais eficientes situam-se entre 2 e 3 (Figuras 4.8, 4.9, e no Anexo A, as Figuras A.9, A.11, A.10 e A.12).

Os gráficos que visualizam o comportamento comparativo das trajetórias dos diversos estimadores em estudo, Figuras 4.10, 4.11 e 4.12, mostram que, para $\xi = 2, 3$, os estimadores que se distinguem no âmbito de E são L_2^{RB} , CH e $GM_{2.5}$, que apresentam trajetórias bastantes estáveis. De notar que, para esses valores de ξ , todos os estimadores apresentam valores de RMSE muito comparáveis. Para $\xi = 1$, são os estimadores L_2^{RB} , MM, $GM_{2.5}$ e CH, que mais se distinguem tendo em conta RMSE.

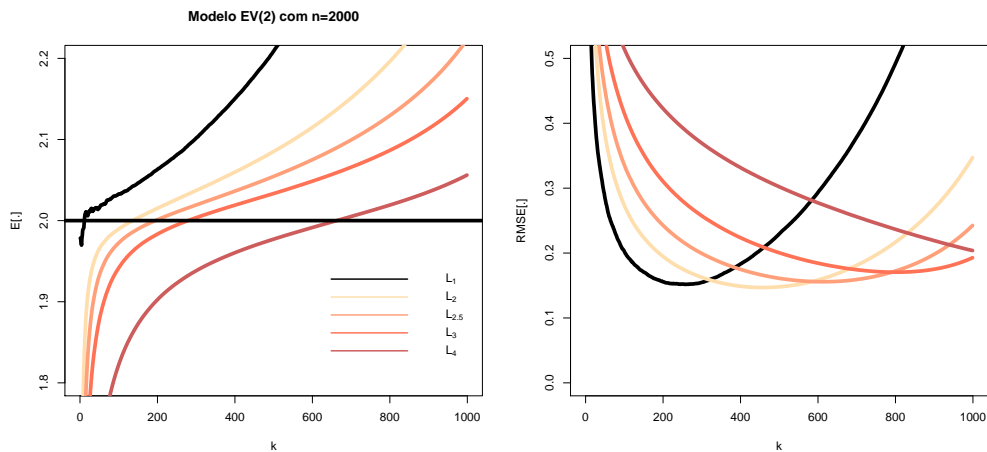


Figura 4.8: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 2, 2.5, 3, 4$.

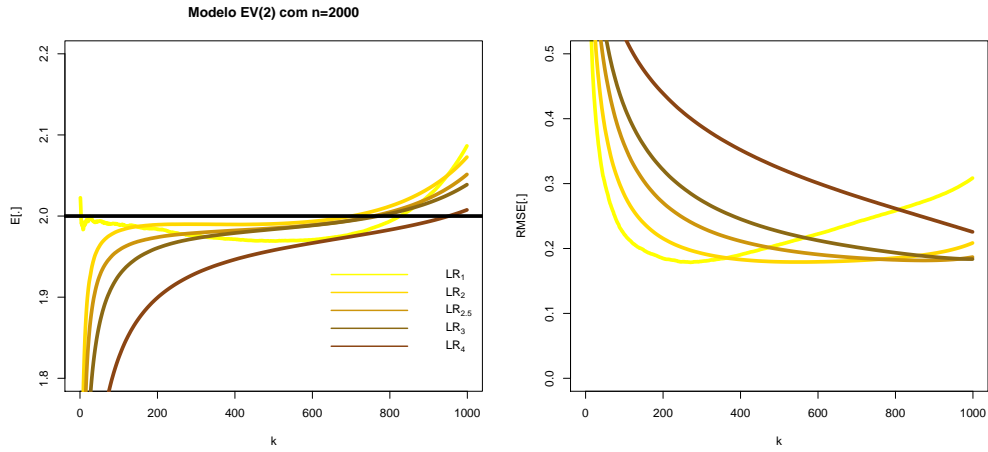


Figura 4.9: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{RB} , $p = 1, 2, 2.5, 3, 4$.

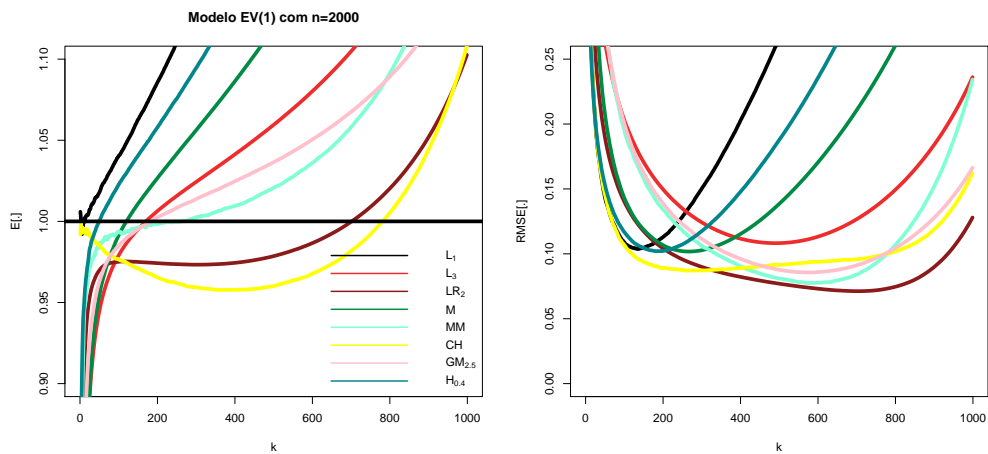


Figura 4.10: Comparação das estimativas simuladas de E e de RMSE.

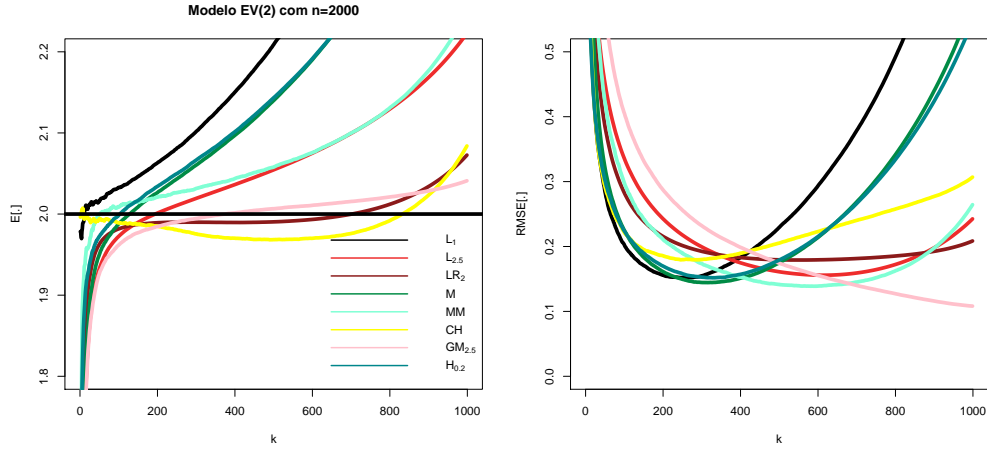


Figura 4.11: Comparação das estimativas simuladas de E e de RMSE.

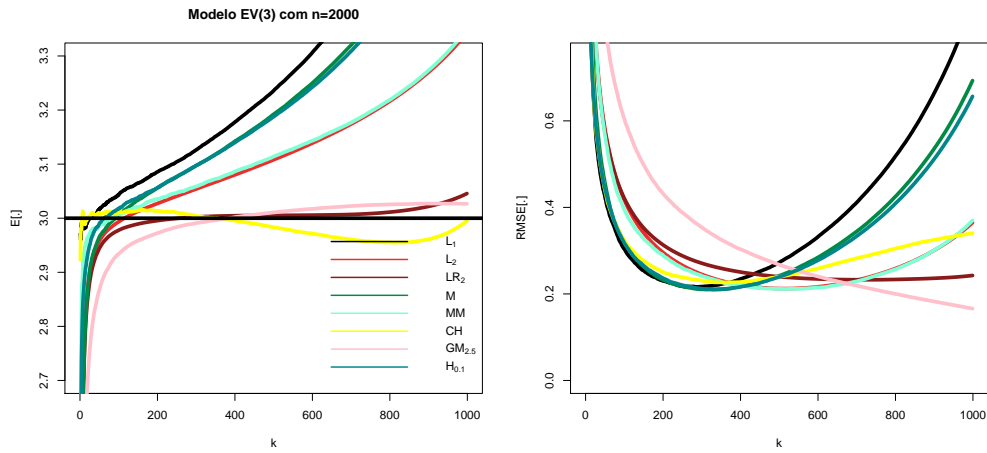


Figura 4.12: Comparação das estimativas simuladas de E e de RMSE.

Modelo $EV(\xi)(\xi = 0.1, 0.5, \rho = -\xi)$

Analisando os gráficos seguintes, referentes ao modelo EV, constata-se que para $\rho > -1$, as trajetórias das estimativas referentes a L e L^{RB} são muito semelhantes, quase se sobrepondo, não parecendo haver grande vantagem nos estimadores que são de viés reduzido face aos estimadores correspondentes que não o são (ver Figuras 4.13, 4.15, 4.14 e 4.16). Observa-se também que o valor de p , que parece otimizar a eficiência dos estimadores $L_p(k)$ e $L_p^{RB}(k)$ é, para $\rho = -0.1$ (ver Figuras 4.13 e 4.15) $p = 30$ e, para $\rho = -0.5$, $p = 8$ (Figuras 4.14 e 4.16).

Os gráficos que visualizam o comportamento comparativo das trajetórias dos di-

versos estimadores em estudo, Figuras 4.17, 4.18 mostram que, para $\rho = -0.1$, os estimadores L_{30} e L_{30}^{RB} , que quase coincidem, juntamente com o estimador GM_5 são os que apresentam trajetórias de RMSE mais baixas. Para $\rho = -0.5$, são os estimadores GM_3 , MM juntamente com os estimadores L_6 e L_6^{RB} que se distinguem.

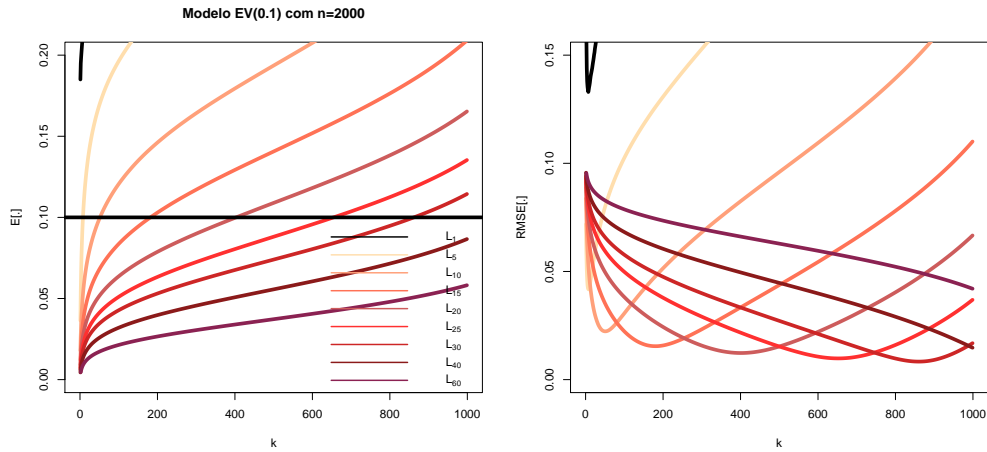


Figura 4.13: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 5(5)30, 40, 60$.

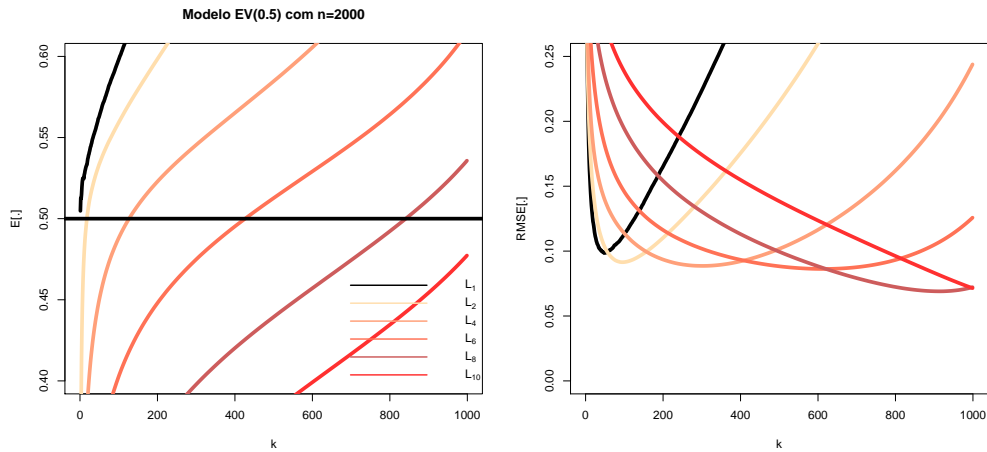


Figura 4.14: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 2(2)10$.

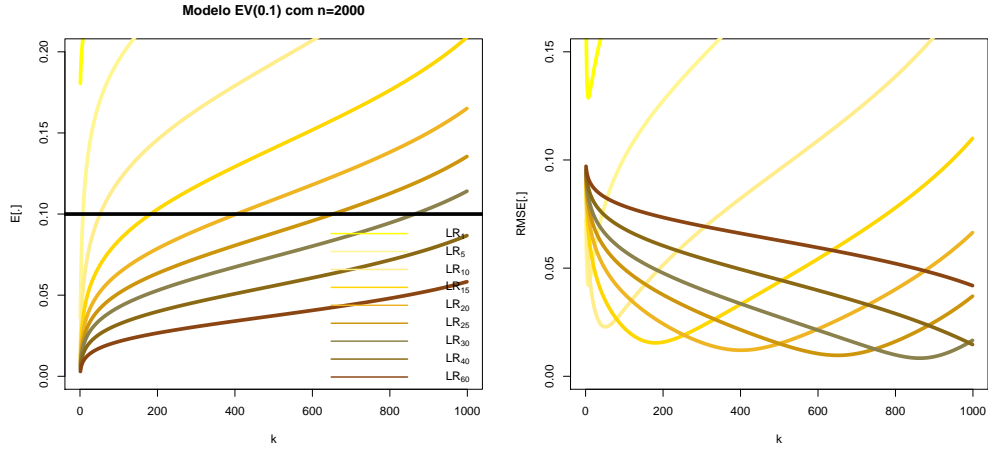


Figura 4.15: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{RB} , $p = 1, 5(5)30, 40, 60$.

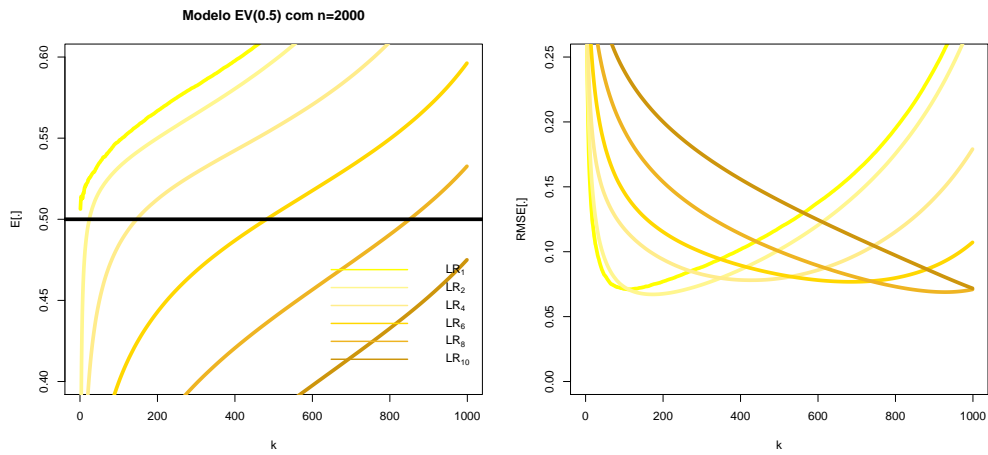


Figura 4.16: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{RB} , $p = 1, 2, (2)10$.

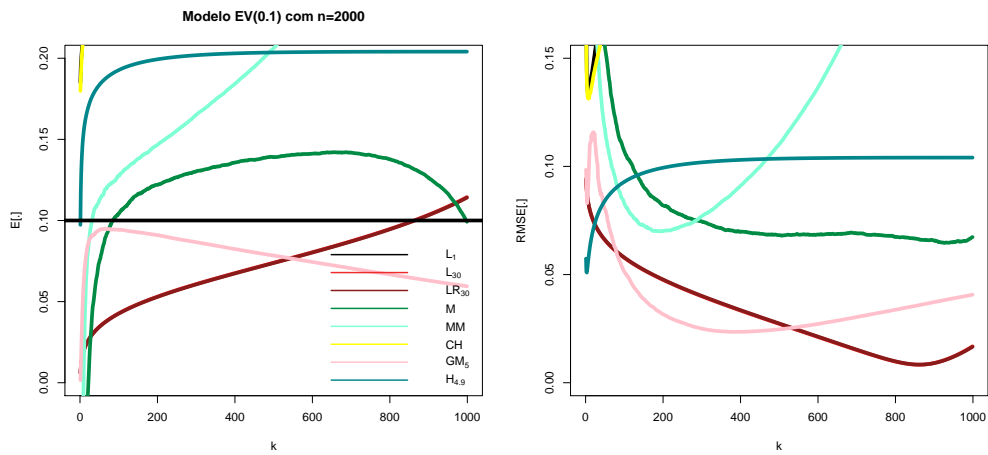


Figura 4.17: Comparação das estimativas simuladas de E e de RMSE.

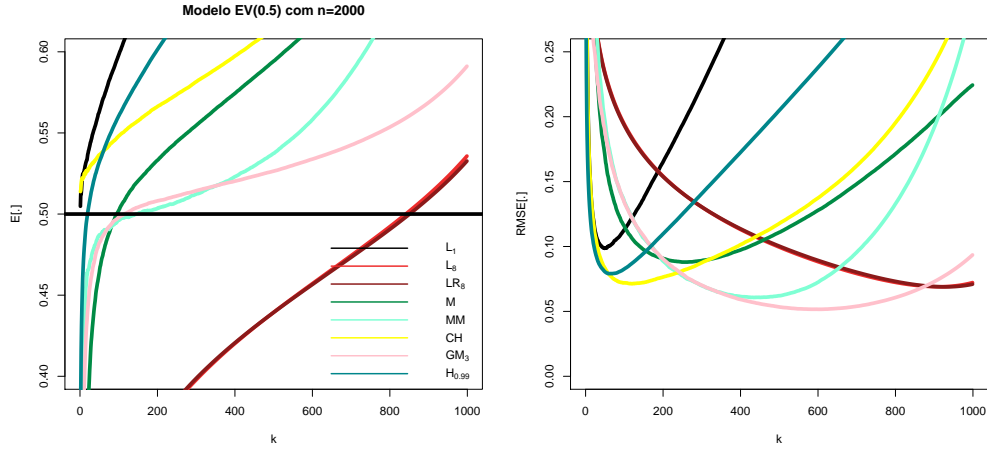


Figura 4.18: Comparação das estimativas simuladas de E e de RMSE.

Modelo Burr($\xi = 0.5, 1, 2, 3; \rho = -0.1$)

Relativamente a este modelo verifica-se que, para $\rho = -0.1$, independentemente do índice de valores extremos ξ , as trajetórias das estimativas de L e L^{RB} são muito semelhantes, quase sobrepondo-se, não parecendo haver grande vantagem nos estimadores que são de viés reduzido face aos estimadores correspondentes que não o são. Observa-se também que, valores de p próximos de 30 parecem melhorar o desempenho dos estimadores $L_p(k)$ (ver Figuras 4.19, e no Anexo A, as Figuras A.13, A.14 e A.15) e dos estimadores $L_p^{\text{RB}}(k)$ (ver Figuras 4.20, e no Anexo A, as Figuras A.16, A.17 e A.18). No gráfico de comparação 4.21 (e no Anexo A, as Figuras A.19, A.20 e A.21) concluí-se que os estimadores L_{30} , L_{30}^{RB} e $GM_{4.5}$ são mais eficientes, sendo que as trajetórias relativas a E são mais estáveis no caso deste último estimador.

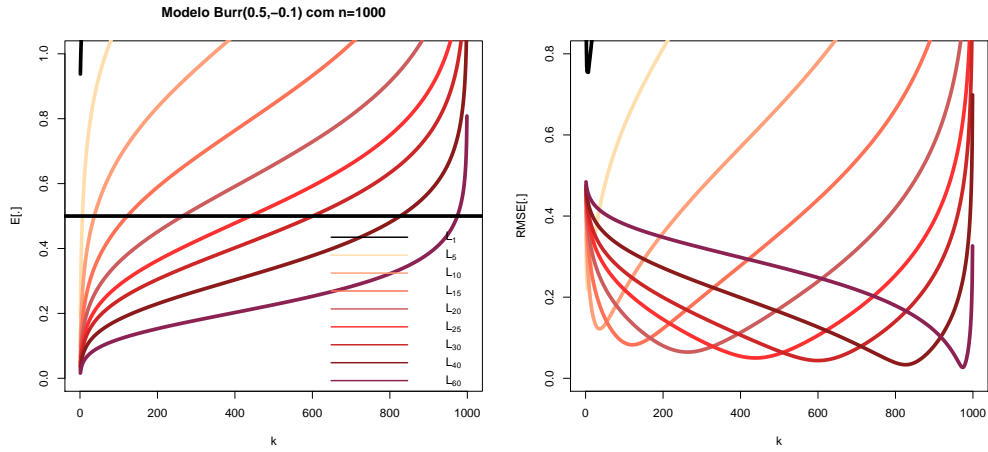


Figura 4.19: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 5(5)30, 40, 60$.

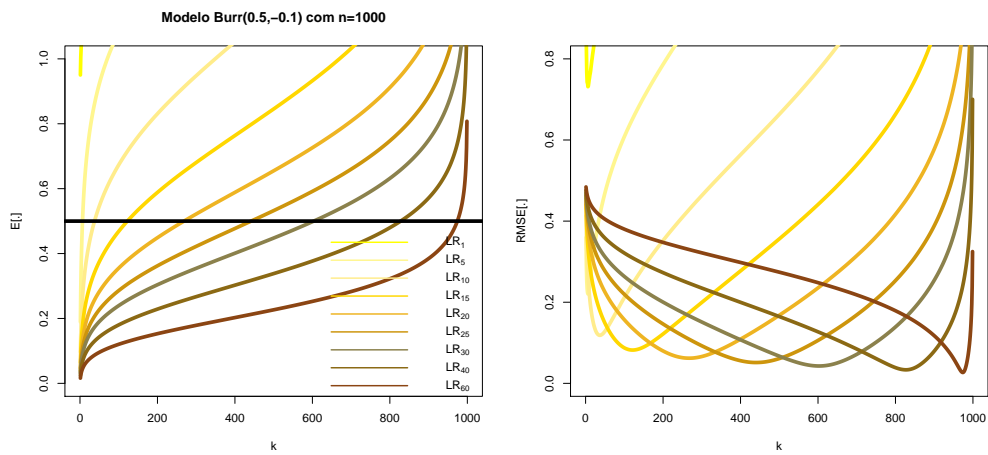


Figura 4.20: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{RB} , $p = 1, 5(5)30, 40, 60$.

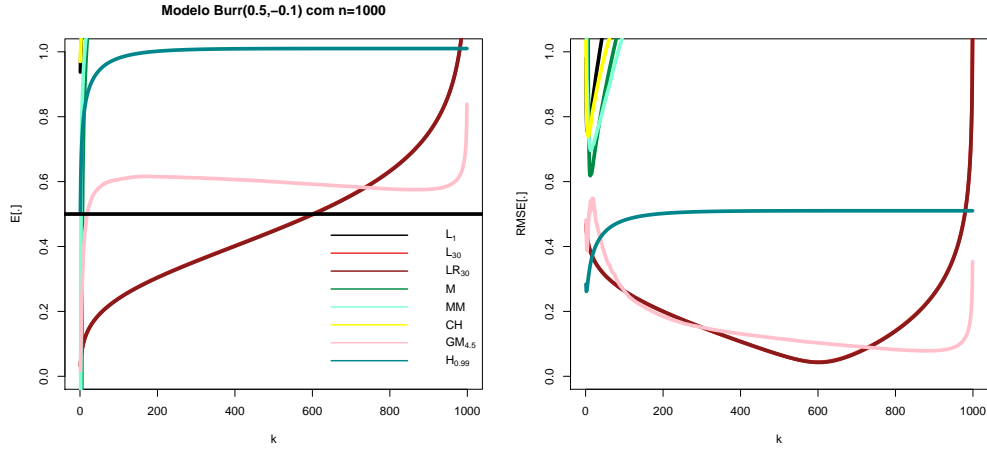


Figura 4.21: Comparação das estimativas simuladas de E e de RMSE.

Modelo Burr($\xi = 0.1, 1, 3; \rho = -0.5$)

Neste caso, independentemente do índice de valores extremos ξ , continua a verificar-se semelhança nas trajetórias de E e de RMSE referentes aos estimadores L e L^{RB} . Um valor de p entre 6 e 8 parece ser adequado para os dois estimadores (ver Figuras 4.22, A.22, A.23 e 4.23, A.24, A.25). Os estimadores com melhor performance parecem ser os estimadores L_8 , L_8^{RB} e GM_3 (ver Figuras 4.24, 4.25 e 4.26).

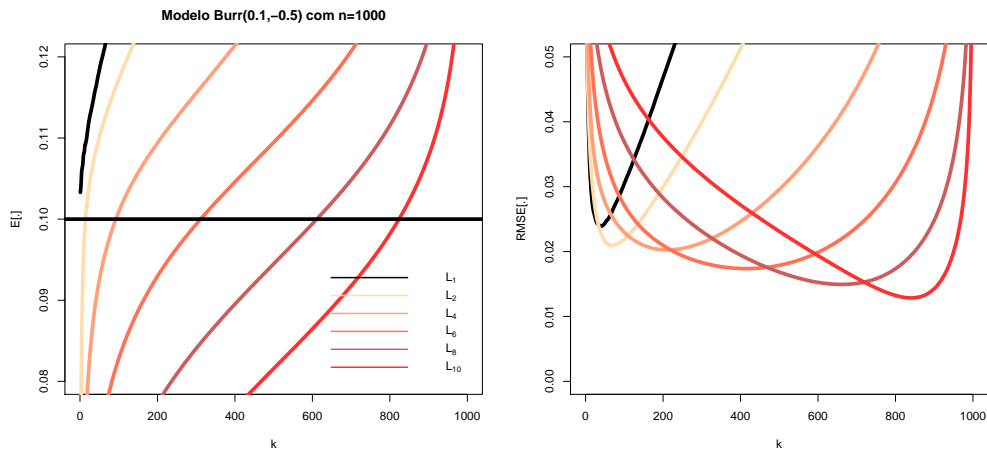


Figura 4.22: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 2(2)10$.

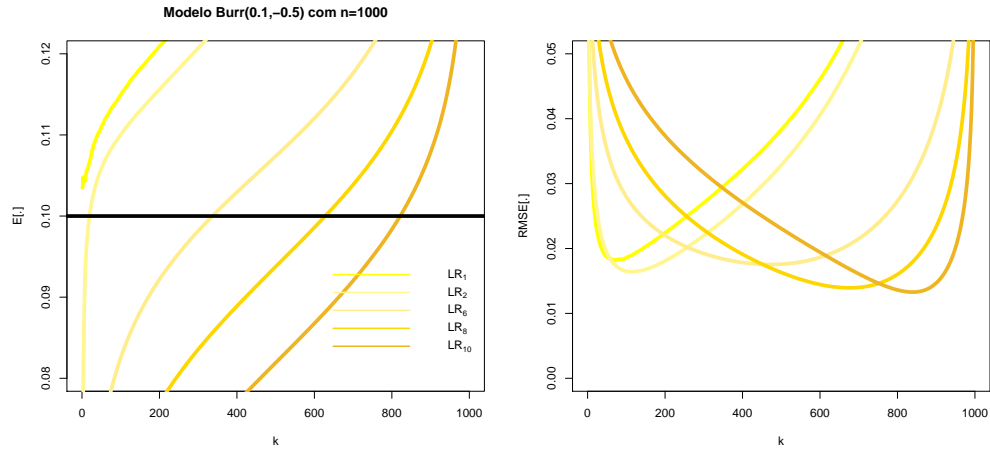


Figura 4.23: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{RB} , $p = 1, 2, 6, 8, 10$.

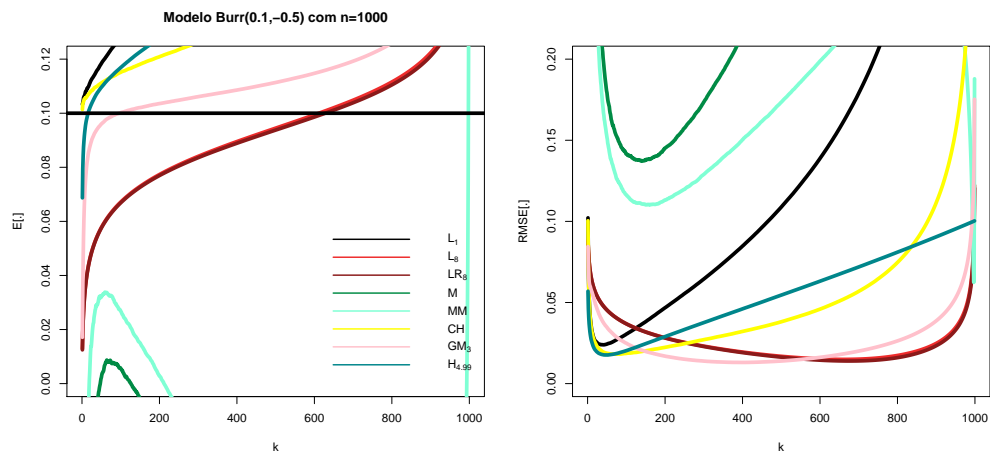


Figura 4.24: Comparação das estimativas simuladas de E e de RMSE.

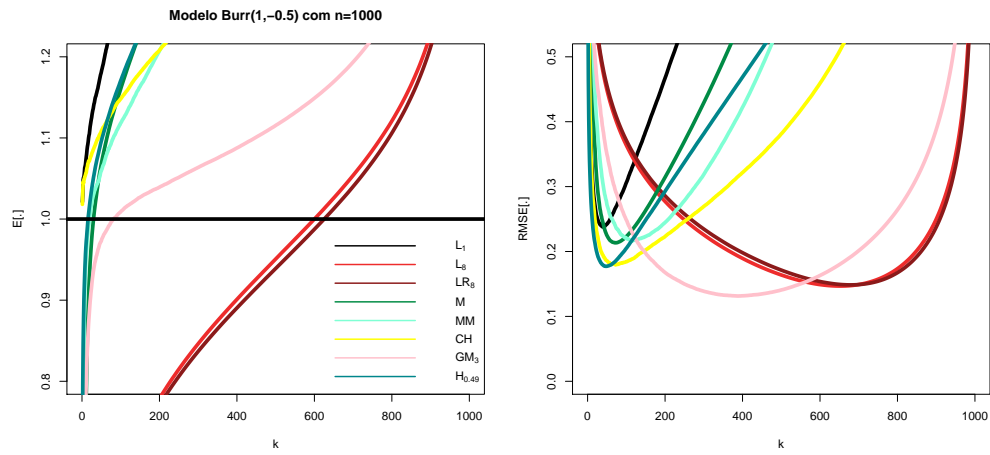


Figura 4.25: Comparação das estimativas simuladas de E e de RMSE.

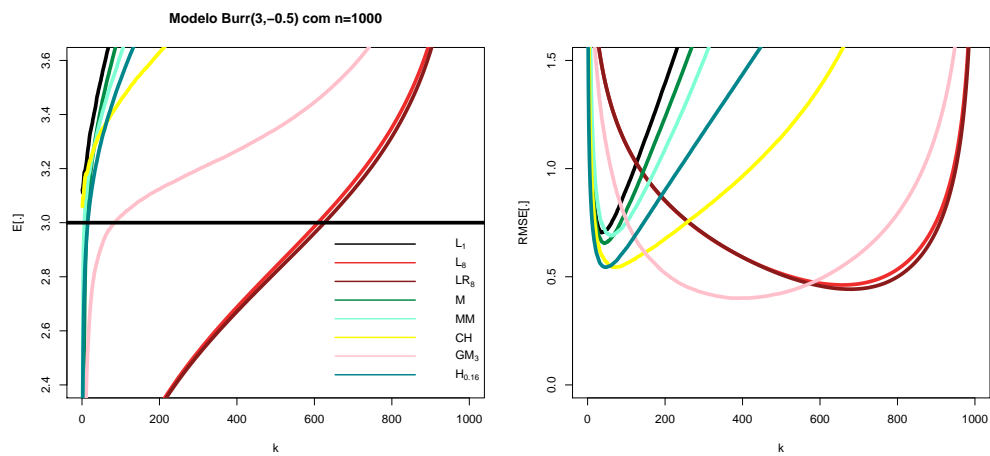


Figura 4.26: Comparação das estimativas simuladas de E e de RMSE.

Modelo Burr($\xi = 1; \rho = -2$)

Nesta situação, valores de p próximos de 2 parecem ser uma escolha adequada para os estimadores L e L^{RB} (ver Figuras 4.27 e 4.28). A Figura 4.29 revela que o estimador $L_2 \equiv \text{GM}_2$ juntamente com o M e o estimador de Hill parecem ser os estimadores mais eficientes.

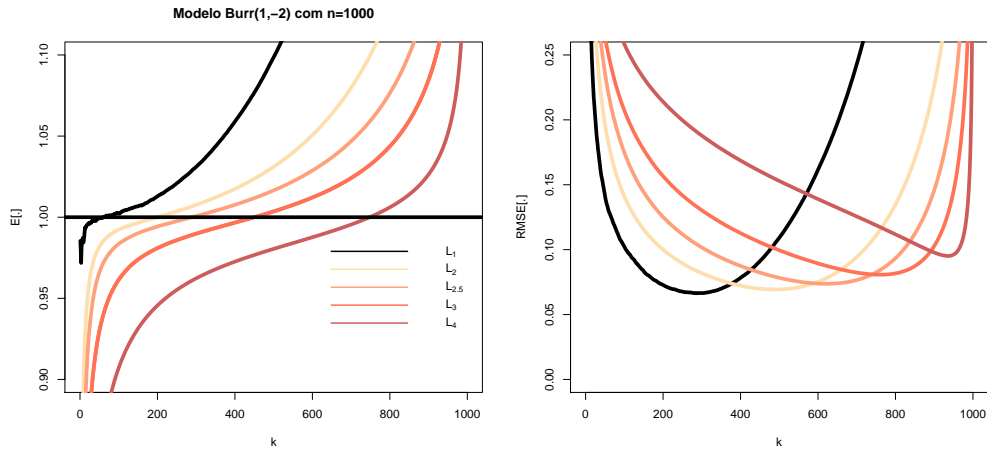


Figura 4.27: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 2, 2.5, 3, 4$.

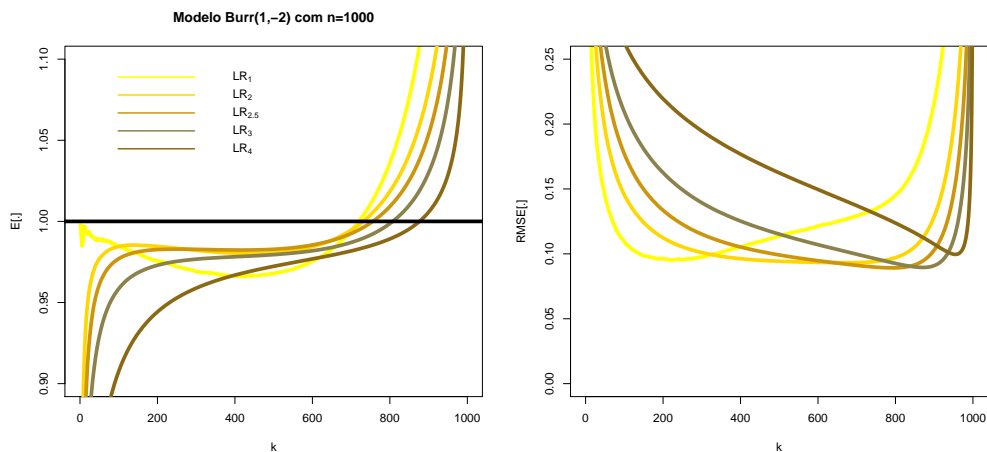


Figura 4.28: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{RB} , $p = 1, 2, 2.5, 3, 4$.

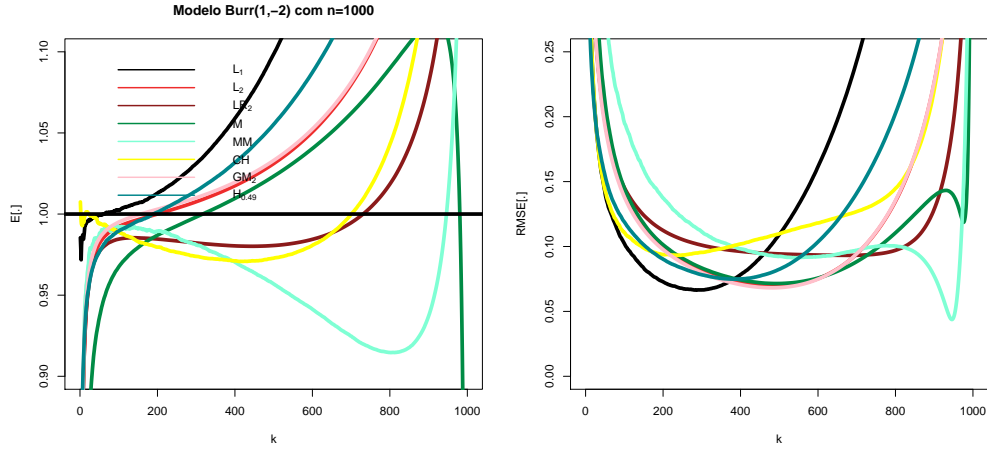


Figura 4.29: Comparação das estimativas simuladas de E e de RMSE.

Modelo Burr($\xi = 0.1, 0.5, 1; \rho = -3$)

Neste caso, independentemente do índice de valores extremos ξ , valores de p próximo de 1.5 parecem ser adequados para os dois estimadores, L e L^{RB} (ver Figuras 4.30, A.27, A.26, e 4.31, A.29, A.28). Os estimadores com melhor performance parecem ser os estimadores $L_{1.5}$, $GM_{1.5}$ e o estimador de Hill (ver Figuras 4.32, 4.33 e 4.33).

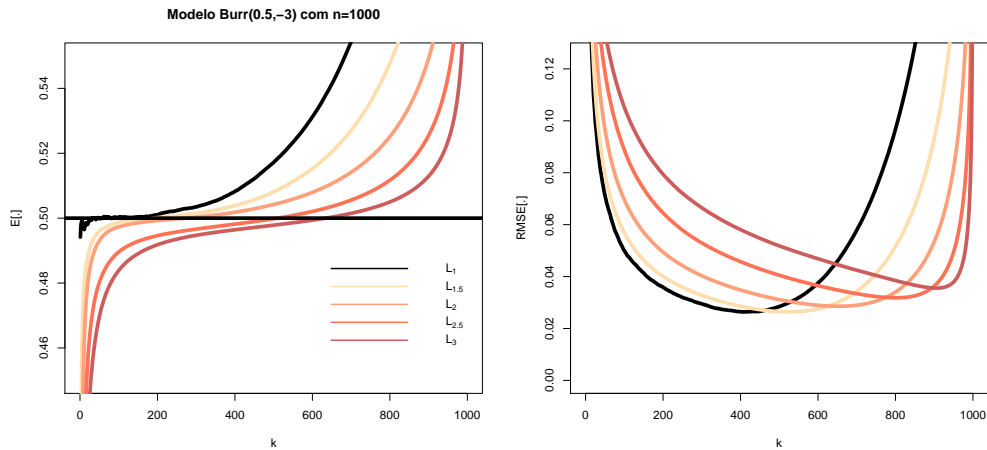


Figura 4.30: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1(0.5)3$.

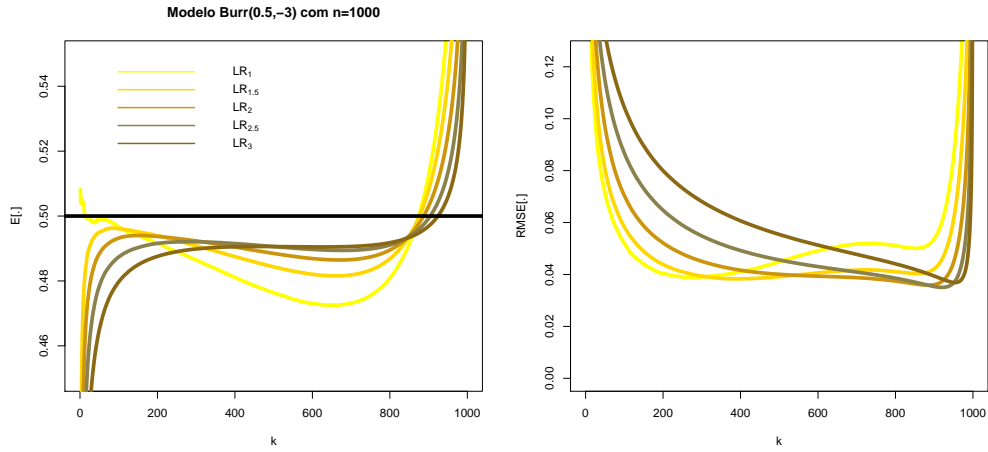


Figura 4.31: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{RB} , $p = 1(0.5)3$.

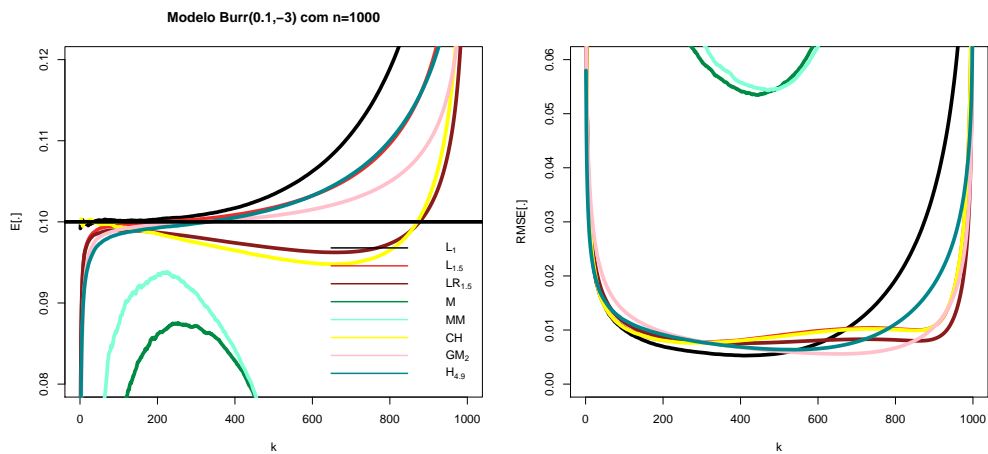


Figura 4.32: Comparação das estimativas simuladas de E e de RMSE.

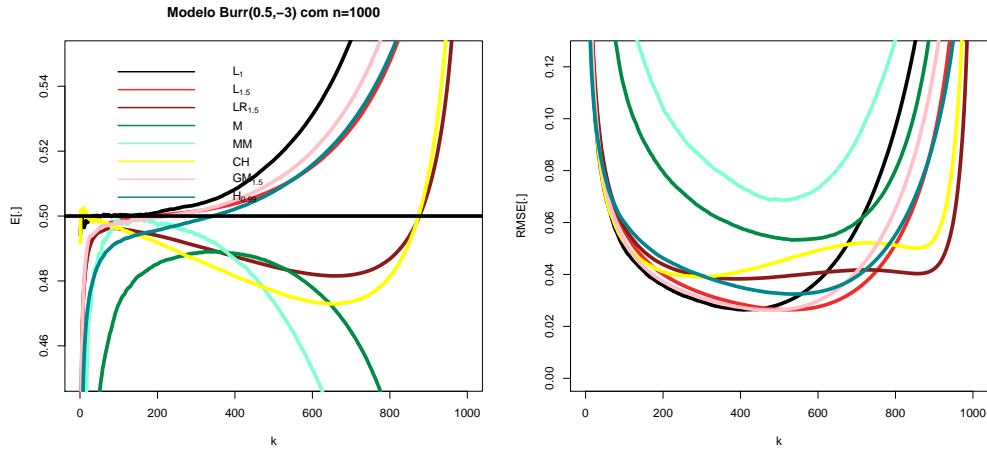


Figura 4.33: Comparação das estimativas simuladas de E e de RMSE.

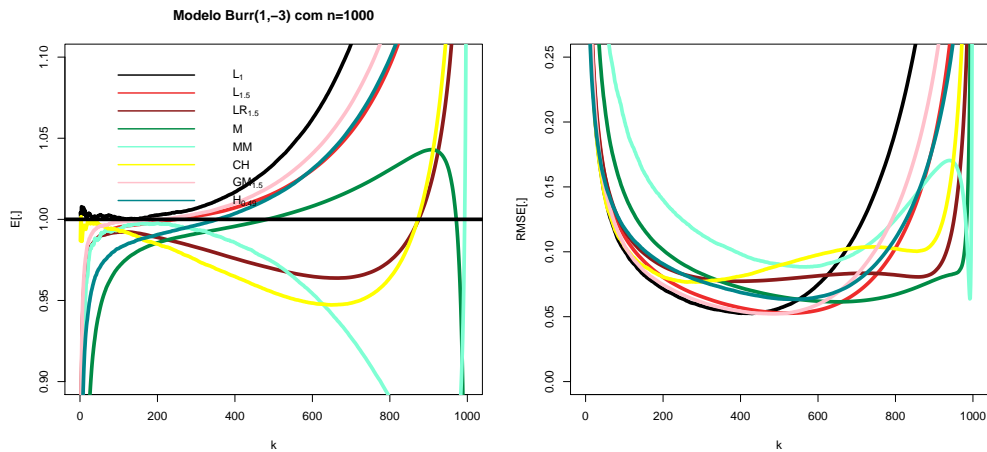


Figura 4.34: Comparação das estimativas simuladas de E e de RMSE.

A Tabela 4.1 mostra, para as diversas situações simuladas e descritas atrás, um quadro resumo com os valores aproximados de p que parecem melhorar o desempenho dos estimadores L e L^{RB} . Tal como esperado assintoticamente, veja-se Figura 3.3(b), verifica-se que estes valores apenas dependem do valor de ρ , decrescendo à medida que o valor de ρ aumenta em valor absoluto, sendo que para $\rho = -0.1$ o valor de p ótimo parece situar-se próximo de 30, enquanto que para $\rho = -3$ verificam-se valores de p próximos de 1.5.

As tabelas 4.2 e 4.3 apresentam respetivamente um resumo dos estimadores que se destacaram segundo o critério de estabilidade de E ou do ponto de vista do RMSE, i.e., que apresentaram valores de RMSE mais baixos numa região mais alargada de valores de k .

Tabela 4.1: Valores de p próximos do p ótimo.

ξ	$\rho = -0.1$		$\rho = -0.5$		$\rho = -1$		$\rho = -2$		$\rho = -3$	
	L	LR	L	LR	L	LR	L	LR	L	LR
0.1	EV : 30	30	Burr : 8	8	Fréchet : 2.5	2	—	—	Burr : 1.5	1.5
0.5	Burr : 30	30	EV : 8	8	Fréchet : 2.5	2	—	—	Burr : 1.5	1.5
1	Burr : 30	30	Burr : 8	8	Fréchet : 2.5	2	Burr : 2	2	Burr : 1.5	1.5
					EV : 3	2				
2	Burr : 30	30	—	—	Fréchet : 2.5	2	—	—	—	—
					EV : 2.5	2				
3	Burr : 30	30	Burr : 8	8	Fréchet : 2.5	2	—	—	—	—
					EV : 2	2				

Valores médios do estimadores-EVI nos níveis ótimos

As tabelas 4.4, 4.5 e 4.6 dizem respeito às simulações efetuadas com as famílias $\text{Burr}_{\xi,\rho}$ e Fréchet_{ξ} ($\rho = -1$). Nelas se apresentam, para $n = 500$ e 1000 , os valores médios simulados nos níveis ótimos (níveis onde os valores de RMSE são mínimos como funções de k) do estimador de Hill e dos estimadores-EVI que se destacaram, ver tabelas 4.2 e 4.3. Informações sobre os intervalos de confiança a 95%, obtidos com base nas 10 réplicas com 5000 runs cada uma, são também providenciadas. Entre os estimadores considerados, aquele que fornece o valor menor do viés ao quadrado é sublinhado e escrito em **bold**.

RMSE e indicadores da eficiência relativa nos níveis ótimos

Nas Tabelas 4.7, 4.8 e 4.9 constam os valores da medida de eficiência relativa, nos níveis ótimos, dos estimadores em estudo, relativamente ao estimador H, $\text{REFF}_{\xi \bullet | \xi^H}$, dada em (4.1). São também fornecidos os correspondentes intervalos de confiança a 95%. Os valores mais elevados dessa medida são sublinhados e escritos em **bold**.

Tabela 4.2: Estimadores com melhor desempenho no conjunto dos modelos estudados, relativamente a E.

ξ	$\rho = -0.1$	$\rho = -0.5$	$\rho = -1$	$\rho = -2$	$\rho = -3$
0.1		Burr : GM ₃ L ₈ L ₈ ^{RB}	Fréchet : CH L ₂ ^{RB}		Burr : L _{1.5} GM ₂ H _{4.9} Hill
0.5	Burr : GM _{4.5}	EV : GM ₃ MM	Fréchet : CH L ₂ ^{RB} GM _{2.5}		Burr : L _{1.5} GM _{1.5} H _{0.99} Hill
1	Burr : GM _{4.5}	Burr : GM ₃	Fréchet : MM CH L ₂ ^{RB} GM _{2.5} EV : MM GM _{2.5} L ₂ ^{RB}	Burr : L ₂ GM ₂ M	Burr : L _{1.5} GM _{1.5} H _{0.49} Hill
2	Burr : GM _{4.5}		Fréchet : CH L ₂ ^{RB} GM _{2.5} EV : L ₂ ^{RB} GM _{2.5} CH		
3	Burr : GM _{4.5}	Burr : GM ₃	Fréchet : CH L ₂ ^{RB} GM _{2.5} EV : L ₂ ^{RB} GM _{2.5} CH		

Tabela 4.3: Estimadores com melhor desempenho no conjunto dos modelos estudados, relativamente a RMSE.

ξ	$\rho = -0.1$	$\rho = -0.5$	$\rho = -1$	$\rho = -2$	$\rho = -3$
0.1	EV : L ₃₀	Burr : L ₈	Fréchet : CH		Burr : L _{1.5}
	LR ₃₀	LR ₈	LR ₂		GM ₂
	GM ₅	GM ₃	L _{2.5}		H _{4.9}
			GM ₂		Hill
					LR _{1.5}
					CH
0.5	Burr : L ₃₀	EV : GM ₃	Fréchet : CH		Burr : L _{1.5}
	LR ₃₀	MM	LR ₂		GM _{1.5}
	GM _{4.5}	LR ₈	GM _{2.5}		H _{0.99}
		L ₈			Hill
1	Burr : L ₃₀	Burr : L ₈	Fréchet : CH	Burr : L ₂	Burr : L _{1.5}
	LR ₃₀	LR ₈	LR ₂	GM ₂	GM _{1.5}
	GM _{4.5}	GM ₃	GM _{2.5}	M	Hill
			MM	Hill	
			EV : LR ₂		
			MM		
			GM _{2.5}		
			CH		
2	Burr : L ₃₀		Fréchet : CH		
	LR ₃₀		LR ₂		
	GM _{4.5}		GM _{2.5}		
			EV : GM _{2.5}		
			MM		
			M		
			H _{0.2}		
			L _{2.5}		
3	Burr : L ₃₀	Burr : L ₈	Fréchet : CH		
	LR ₃₀	LR ₈	LR ₂		
	GM _{4.5}	GM ₃	GM _{2.5}		
			EV : GM _{2.5}		
			MM		
			L ₂		

Tabela 4.4: Valores médios simulados nos níveis ótimos e os correspondentes intervalos de confiança a 95%, relativos à família $\text{Burr}_{\xi,-0.5}$.

n	500	1000
$\text{Burr}_{0.5,-0.5}$		
$L_1 \equiv H$	0.594 ± 0.0015	0.579 ± 0.0008
L_4	0.548 ± 0.0007	0.549 ± 0.0009
L_6	<u>0.522</u> ± 0.0008	0.527 ± 0.0010
L_8	0.560 ± 0.0006	0.523 ± 0.0006
L_{10}	0.629 ± 0.0005	<u>0.515</u> ± 0.0005
MM	0.554 ± 0.0007	0.544 ± 0.0005
GM_3	0.551 ± 0.0007	0.544 ± 0.0006
$\text{Burr}_{1,-0.5}$		
$L_1 \equiv H$	1.192 ± 0.0025	1.151 ± 0.0013
L_4	1.097 ± 0.0023	1.100 ± 0.0016
L_6	<u>1.045</u> ± 0.002	1.055 ± 0.0011
L_8	1.122 ± 0.0018	1.046 ± 0.0017
L_{10}	1.258 ± 0.0010	<u>1.029</u> ± 0.0010
GM_3	1.104 ± 0.0009	1.087 ± 0.0010

Tabela 4.5: Valores médios simulados nos níveis ótimos e os correspondentes intervalos de confiança a 95%, relativos à família $\text{Fréchet}_\xi(\rho = -1)$.

n	500	1000
Fréchet _{0.5}		
$L_1 \equiv H$	0.532 ± 0.0005	0.525 ± 0.0003
$L_{1.5}$	0.524 ± 0.0004	0.523 ± 0.0004
L_2	0.529 ± 0.0007	0.523 ± 0.0002
$L_1^{\text{RB}} \equiv CH$	<u>0.499</u> ± 0.0009	<u>0.500</u> ± 0.0003
$L_{1.5}^{\text{RB}}$	0.503 ± 0.0009	0.502 ± 0.0003
$GM_{2.5}$	0.523 ± 0.0011	0.521 ± 0.0002
Fréchet ₁		
$L_1 \equiv H$	1.063 ± 0.0007	1.050 ± 0.0007
$L_{1.5}$	1.058 ± 0.0013	1.046 ± 0.0011
L_2	1.058 ± 0.0002	1.047 ± 0.0004
$L_1^{\text{RB}} \equiv CH$	<u>0.998</u> ± 0.0000	<u>0.999</u> ± 0.0000
$L_{1.5}^{\text{RB}}$	1.007 ± 0.0023	0.976 ± 0.0008
$GM_{2.5}$	1.050 ± 0.0015	1.042 ± 0.0004

Tabela 4.6: Valores médios simulados nos níveis ótimos e os correspondentes intervalos de confiança a 95%, relativos à família $\text{Burr}_{\xi,-2}$.

n	500	1000
$\text{Burr}_{0.5,-2}$		
$L_1 \equiv H$	0.520 ± 0.0004	0.525 ± 0.0003
$L_{1.5}$	0.514 ± 0.0003	0.514 ± 0.0003
$L_2 \equiv \text{GM}_2$	0.518 ± 0.0004	0.515 ± 0.0004
$L_{2.5}$	0.520 ± 0.0007	0.516 ± 0.0004
L_3	0.522 ± 0.0005	0.518 ± 0.0004
$L_1^{\text{RB}} \equiv \text{CH}$	0.481 ± 0.0007	0.488 ± 0.0000
$L_{1.5}^{\text{RB}}$	0.481 ± 0.0006	0.488 ± 0.0006
L_2^{RB}	<u>0.499</u> ± 0.0004	0.495 ± 0.0004
$L_{2.5}^{\text{RB}}$	0.506 ± 0.0007	<u>0.504</u> ± 0.0004
L_3^{RB}	0.5113 ± 0.0004	0.509 ± 0.0004
M	0.468 ± 0.0008	0.476 ± 0.0006
$\text{GM}_{2.5}$	0.509 ± 0.0002	0.506 ± 0.0003
$\text{Burr}_{1,-2}$		
$L_1 \equiv H$	1.038 ± 0.0006	1.029 ± 0.0006
$L_{1.5}$	1.035 ± 0.0006	1.027 ± 0.0004
$L_2 \equiv \text{GM}_2$	1.037 ± 0.0008	1.029 ± 0.0007
$L_{2.5}$	1.040 ± 0.0009	1.032 ± 0.0006
L_3	1.043 ± 0.0010	1.035 ± 0.0010
$L_1^{\text{RB}} \equiv \text{CH}$	0.962 ± 0.0009	0.976 ± 0.0090
$L_{1.5}^{\text{RB}}$	0.963 ± 0.0012	0.977 ± 0.0007
L_2^{RB}	<u>0.999</u> ± 0.0008	0.990 ± 0.0009
$L_{2.5}^{\text{RB}}$	1.011 ± 0.0011	<u>1.007</u> ± 0.0005
L_3^{RB}	1.022 ± 0.0007	1.017 ± 0.0009
M	1.026 ± 0.0010	1.022 ± 0.0006
$\text{GM}_{2.5}$	1.018 ± 0.0007	1.011 ± 0.0004

Tabela 4.7: RMSE simulado de H (primeira linha) no nível ótimo e indicadores REFF juntamente com os correspondentes intervalos de confiança a 95%, relativos à família $\text{Burr}_{\xi,-0.5}$.

n	500	1000
$\text{Burr}_{0.5,-0.5}$		
RMSE ₀ (H)	0.145 ± 0.0012	0.118 ± 0.0008
L ₄	1.286 ± 0.0210	1.188 ± 0.0128
L ₆	1.535 ± 0.0269	1.320 ± 0.0206
L ₈	1.560 ± 0.0167	1.536 ± 0.0186
L ₁₀	1.034 ± 0.0049	<u>1.808</u> ± 0.0247
MM	1.408 ± 0.0085	1.450 ± 0.0085
GM ₃	<u>1.825</u> ± 0.0121	1.766 ± 0.0162
$\text{Burr}_{1,-0.5}$		
RMSE ₀ (H)	0.290 ± 0.0022	0.235 ± 0.0008
L ₄	1.273 ± 0.0214	1.169 ± 0.0113
L ₆	1.507 ± 0.0317	1.312 ± 0.0190
L ₈	1.547 ± 0.0217	1.531 ± 0.0322
L ₁₀	1.034 ± 0.0061	<u>1.771</u> ± 0.0237
GM ₃	<u>1.807</u> ± 0.0170	1.769 ± 0.0163

Tabela 4.8: RMSE simulado de H (primeira linha) no nível ótimo e indicadores REFF juntamente com os correspondentes intervalos de confiança a 95%, relativos à família $\text{Fréchet}_\xi(\rho = -1)$.

n	500	1000
Fréchet _{0.5}		
RMSE ₀ (H)	0.059 ± 0.0004	0.046 ± 0.0002
L _{1.5}	1.049 ± 0.0062	1.060 ± 0.0059
L ₂	1.042 ± 0.0103	1.038 ± 0.0070
L ₁ ^{RB} ≡ CH	1.117 ± 0.0190	1.148 ± 0.0053
L _{1.5} ^{RB}	1.123 ± 0.0118	1.160 ± 0.0103
GM _{2.5}	<u>1.300</u> ± 0.0190	<u>1.272</u> ± 0.0122
Fréchet ₁		
RMSE ₀ (H)	0.117 ± 0.0002	0.092 ± 0.0008
L _{1.5}	1.071 ± 0.0135	1.063 ± 0.0098
L ₂	1.050 ± 0.0137	1,039 ± 0.0079
L ₁ ^{RB} ≡ CH	1.123 ± 0.0177	1.155 ± 0.0113
L _{1.5} ^{RB}	1.144 ± 0.0982	0,985 ± 0.0044
GM _{2.5}	<u>1.311</u> ± 0.0215	<u>1.280</u> ± 0.0068

Tabela 4.9: RMSE simulado de H (primeira linha) no nível ótimo e indicadores REFF juntamente com os correspondentes intervalos de confiança a 95%, relativos à família $\text{Burr}_{\xi,-2}$.

n	500	1000
$\text{Burr}_{0.5,-2}$		
$\text{RMSE}_0(\text{H})$	0.044 ± 0.0003	0.033 ± 0.0002
$L_{1.5}$	1.037 ± 0.0085	1.027 ± 0.0064
$L_2 \equiv \text{GM}_2$	0.988 ± 0.0061	0.971 ± 0.0084
$L_{2.5}$	0.919 ± 0.0072	0.894 ± 0.0100
L_3	0.843 ± 0.0099	0.814 ± 0.0127
$L_1^{\text{RB}} \equiv \text{CH}$	0.734 ± 0.0055	0.701 ± 0.0048
$L_{1.5}^{\text{RB}}$	0.747 ± 0.0066	0.712 ± 0.0037
L_2^{RB}	0.808 ± 0.0059	0.715 ± 0.0039
$L_{2.5}^{\text{RB}}$	0.830 ± 0.0032	0.743 ± 0.0021
L_3^{RB}	0.812 ± 0.0059	0.744 ± 0.0040
M	0.528 ± 0.0021	0.550 ± 0.0034
$\text{GM}_{2.5}$	<u>1.391</u> ± 0.0165	<u>1.430</u> ± 0.0151
$\text{Burr}_{1,-2}$		
$\text{RMSE}_0(\text{H})$	0.089 ± 0.0009	0.066 ± 0.0005
$L_{1.5}$	1.059 ± 0.0067	1.031 ± 0.0086
$L_2 \equiv \text{GM}_2$	1.008 ± 0.0071	0.969 ± 0.0060
$L_{2.5}$	0.936 ± 0.0009	0.896 ± 0.0088
L_3	0.869 ± 0.0010	0.831 ± 0.0117
$L_1^{\text{RB}} \equiv \text{CH}$	0.749 ± 0.0046	0.709 ± 0.0037
$L_{1.5}^{\text{RB}}$	0.768 ± 0.0042	0.721 ± 0.0041
L_2^{RB}	0.821 ± 0.0041	0.729 ± 0.0040
$L_{2.5}^{\text{RB}}$	0.846 ± 0.0067	0.756 ± 0.0043
L_3^{RB}	0.828 ± 0.0049	0.756 ± 0.0034
M	0.923 ± 0.0060	0.930 ± 0.0045
$\text{GM}_{2.5}$	<u>1.417</u> ± 0.0128	<u>1.457</u> ± 0.0107

Alguns comentários e conclusões

1. Verifica-se que os valores de p que parecem melhorar o desempenho dos estimadores L e L^{RB} dependem fortemente do valor de ρ . Constata-se que, esses valores são, para $\rho > -1$, consideravelmente maiores do que aqueles que se verificam para $\rho \leq -1$. Especificamente, para $\rho = -0.1$, estes valores situam-se próximo de 30 e para $\rho = -0.5$, próximo de 8. Já quando o $\rho \leq -1$, os valores de p que tornam os estimadores $L_p(k)$ e $L_p^{RB}(k)$ mais eficientes situam-se entre 1.5 e 2.5.
2. Para $\rho > -1$ verifica-se que, independentemente do índice de valores extremos ξ , as trajetórias das estimativas de L e L^{RB} são muito semelhantes, quase sobrepondo-se, não parecendo haver grande vantagem nos estimadores que são de viés reduzido face aos estimadores correspondentes que não o são.
3. A classe de estimadores L , para valores de p adequados, apresenta uma estabilidade em torno do verdadeiro valor do parâmetro superior quando $\rho < -1$.
4. Quando $\rho = -1$, e para o modelo Fréchet, independentemente do valor de ξ , as trajetórias das estimativas referentes a L^{RB} , para valores de p adequados, exibem uma estabilidade elevada em torno do verdadeiro valor do parâmetro. O mesmo se verifica, para o modelo EV, quando $\xi = 1, 2, 3$.
5. Para $\rho \leq -1$, as trajetórias das estimativas referentes a L^{RB} são sempre mais estáveis do que as trajetórias dos estimadores L . No entanto, para o modelo Burr, verifica-se que a estabilidade das estimativas de L^{RB} se realiza em torno de um valor que é inferior ao valor do parâmetro.
6. De uma maneira geral, as classes de estimadores que parecem revelar uma melhor performance são, para valores de p ótimos:
 - (a) as classes de estimadores L e GM, quando $\rho > -1$;
 - (b) as classes de estimadores L^{RB} e GM, quando $\rho = -1$;
 - (c) as classes de estimadores L e GM, quando $\rho < -1$, apesar da eficiência não ser muito maior quando comparado com o estimador H.

Capítulo 5

Casos de Estudo - Aplicação a Dados Reais

Neste ponto ilustram-se as abordagens paramétrica e semi-paramétrica na estimação do parâmetro EVI.

Os trabalhos de Penalva *et al.* (2013) e Penalva *et al.* (2014b) apresentam o conjunto de passos necessários para realizar uma análise de valores extremos, explorando os *packages* e funções existentes no R para realizar tal análise.

O *software* R (R Core Team, 2013) linguagem de código aberto para computação estatística e tratamento de dados, possui vários *packages* estatísticos e permite a implementação pelos utilizadores de *packages* adicionais. Para a modelação e estimação em EVT, quer na abordagem paramétrica quer na semi-paramétrica, podem-se referir alguns *packages* que o *software* R possui: `evd`; `ismev`; `evir`; `fExtremes`; `POT`; `evdbayes`; `copula`; `SpatialExtremes`.

Nesses trabalhos, apenas a abordagem paramétrica foi abordada. Especificamente, no trabalho de Penalva *et al.* (2013), foi considerada a metodologia de BM, envolvendo o método de estimação ML e o método de verosimilhança de perfil. Estas metodologias foram aplicadas a dois conjuntos de dados, que podem ser obtidos no R, no *package* `evd`: 1) velocidades máximas anuais do vento, em quilómetros por hora, registadas em Lisboa, no período de 1941 a 1970; 2) descargas máximas anuais, me-

didadas em pés cúbicos por segundo, referentes às inundações verificadas a norte do rio Saskachevan em Edmonton, durante um período de 47 anos. No trabalho de Penalva *et al.* (2014b), foram também incluídas a abordagem POT e a abordagem k -LO. Os métodos de estimação utilizados foram os dois métodos de estimação referidos atrás e o método PWM. De referir, que neste último trabalho, também os níveis de retorno foram estimados. Foram usados os *packages* *evd*, *isnev*, *fExtremes* e *evir*. Os dados que serviram para ilustrar as diversas metodologias representam a precipitação diária acumulada, em milímetros, numa localidade no sudoeste de Inglaterra. Estes dados estão disponíveis no *package* *isnev*.

A aplicação da abordagem semi-paramétrica pode ser vista nos trabalhos de Penalva *et al.* (2014a) e Penalva *et al.* (2016b). Estes trabalhos tiveram como objetivo comparar as abordagens paramétrica e semi-paramétrica na estimação de EVI, recorrendo ao *software* R. Na abordagem paramétrica foram usados os mesmos *packages* já referidos, enquanto que, para obter os resultados da estimação semi-paramétrica, foram desenvolvidas e implementadas, em linguagem R, todas as necessárias funções. De referir ainda que neste último trabalho, foram ainda incluídos o estudo e aplicação de testes estatísticos que permitem validar a condição de primeira ordem de valores extremos e aqueles que determinam qual é o domínio de atração mais adequado aos dados em análise. As metodologias foram aplicadas às descargas médias diárias, em metros cúbicos por segundo, do rio Paiva, medidas na estação hidrométrica de Fragas da Torre, no período entre 1946/47 e 1995/06.

Nas secções seguintes são apresentados dois casos de estudo de aplicação dos métodos descritos ao longo do trabalho. Na Secção 5.1 apresenta-se um resumo dos resultados obtidos em Penalva *et al.* (2016b). Na Secção 5.2 ilustra-se a aplicação dos estimadores semi-paramétricos em estudo a um conjunto de dados de áreas ardidas em Portugal.

5.1 Caso de estudo: descargas médias diárias no rio Paiva

Os dados analisados consistem nas descargas médias diárias medidas na estação hidrométrica de Fragas da Torre no rio Paiva. A nascente deste rio localiza-se na serra de Leomil, a norte de Portugal. Este rio é um afluente do rio Douro, com uma bacia hidrográfica de aproximadamente 700 km^2 . O estudo do fluxo deste rio é extremamente importante uma vez que é uma das alternativas ao rio Douro como fonte de abastecimento de água para a região sul do Porto. De uma forma mais precisa, o conjunto de dados em estudo representa as descargas médias diárias (m^3/s) observadas de 1 de Outubro de 1946 a 30 de Setembro de 2006, recolhidas a partir do "SNIRH: Sistema Nacional de Informação dos Recursos Hídricos".

Para obtenção dos resultados, recorreu-se aos *packages* do *software* R.

5.1.1 Breve análise preliminar dos dados

Após algumas análises gráficas prévias, mostrando a ocorrência dos valores extremos nos diferentes meses, conselhos de hidrólogos e levando em consideração um trabalho anterior que considerou alguns anos iniciais desses dados, Gomes (1993), apenas os meses de novembro até abril foram utilizados em cada ano. Tem-se assim um total de 10860 observações.

Os principais resultados de uma análise gráfica e descritiva preliminar são mostrados na Tabela 5.1 e na Figura 5.1.

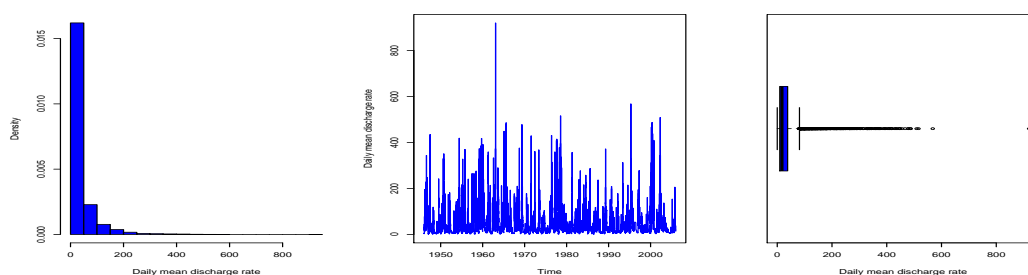


Figura 5.1: Histograma (esquerda); cronograma (centro) e caixa-com-bigodes (direita).

Tabela 5.1: Estatísticas descritivas básicas.

n	Min	1º Quart.	Mediana	Média	3º Quart.	Max	Desv. Padrão.	Assim.	Kurt.
10860	0	9.20	17.30	34.83	38.00	920.00	50.92	4.15	27.31

O histograma, a caixa com bigodes, as estatísticas descritivas, particularmente o coeficiente de assimetria, parecem indicar uma cauda direita pesada.

5.1.2 Teste à condição de valores extremos

Em todos os contextos de estimação em EVT, quer na abordagem paramétrica, como também na semi-paramétrica, é admitido que a f.d. F subjacente aos dados pertence ao $\mathcal{D}_{\mathcal{M}}(EV_{\xi})$, para um valor apropriado de ξ , ou que pertence a um sub-domínio específico de $\mathcal{D}_{\mathcal{M}}(EV_{\xi})$. Antes de se realizar uma aplicação prática em valores extremos, é importante averiguar se essa condição é razoável para o conjunto de dados em estudo. Assim deve-se testar:

$$H_0 : F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi}) \text{ para um } \xi \in \mathbb{R}. \quad (5.1)$$

Dietrich *et al.* (2002) propuseram os testes E , PE (se se admitir $\xi \geq 0$) e Drees *et al.* (2006) o teste T (admitindo $\xi > -1/2$).

Para $\eta > 0$, Dietrich *et al.* (2002) introduziram as seguintes estatísticas de teste:

$$E_n := k \int_0^1 \left(\frac{\log X_{n-[kt],n} - \log X_{n-k,n}}{\hat{\xi}_+} - \frac{t^{-\hat{\xi}_-} - 1}{\hat{\xi}_-} (1 - \hat{\xi}_-) \right)^2 t^{\eta} dt, \quad (5.2)$$

$$PE_n := k \int_0^1 \left(\frac{\log X_{n-[kt],n} - \log X_{n-k,n}}{\hat{\xi}_+} + \log t \right)^2 t^{\eta} dt, \quad (5.3)$$

onde as estimativas para $\xi_+ := \max(0, \xi)$ e $\xi_- := \min(0, \xi)$ são obtidas através dos estimadores dos momentos em Dekkers *et al.* (1989), e $k \equiv k(n)$ é uma sucessão intermédia de inteiros satisfazendo, $k \equiv k_n \rightarrow \infty$, $k/n \rightarrow 0$ e $k^{1/2} A(n/k) \rightarrow 0$, quando $n \rightarrow \infty$.

Drees *et al.* (2006) propuseram a estatística de teste,

$$T_n := k \int_0^1 \left(\frac{n}{k} \bar{F}_n \left(\hat{a}_{n/k} \frac{x^{-\hat{\xi}} - 1}{\hat{\xi}} + \hat{b}_{n/k} \right) - x \right)^2 x^{\eta-2} dx, \quad (5.4)$$

onde $\bar{F}_n = 1 - F_n$ com F_n , a f.d. empírica. Recomenda-se o uso de estimadores ML para ξ e a e $\hat{b}_{n/k} := X_{n-k:n}$.

Hüsler and Li (2006) realizaram um estudo de comparação desses testes, providenciando algumas correções às correspondentes tabelas de valores críticos, e também realizaram um extenso estudo de simulação com diretrizes para obter o valor de η . Estes autores desenvolveram um *package*, no *software* R, o `TestEVC1d`, constituído por algoritmos que permitem a obtenção das trajetórias das estatísticas de teste referidas em cima.

Para o conjunto de dados em análise, testou-se a hipótese em (5.1), através da estatística E_n , em (5.2). A Figura 5.2 mostra a trajetória da estatística E_n e observa-se que na sua maioria os seus valores são menores que os correspondentes quantis de probabilidade 0.95. Assim uma vez que a trajetória amostral da estatística de teste se encontra na sua maioria fora da região de rejeição, não existe razão para rejeitar a hipótese nula, em (5.1).

5.1.3 A metodologia BM

Neste âmbito, consideraram-se os anos como blocos de observações e escolheram-se como valores extremos os valores máximos em cada ano, constituindo um conjunto de 60 observações. Com estes dados obteve-se um coeficiente de assimetria= 0.998 e Kurtose=2.265 e as análises gráficas são mostradas na Figura 5.3.

O histograma, a caixa-com-bigodes e o coeficiente de assimetria indicam uma assimetria positiva moderada. A partir do gráfico da função de autocorrelação parcial (ACF), parece ser razoável assumir que os dados não são correlacionados. Assim pode-se ajustar a distribuição EV_ξ aos máximos de cada ano. Na tabela 5.2

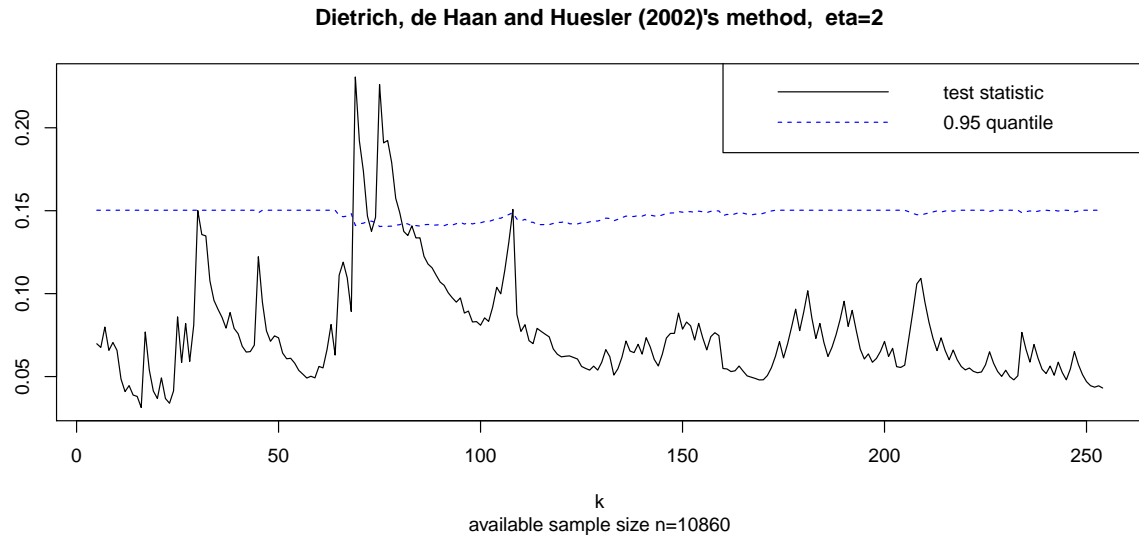


Figura 5.2: Trajetória amostral da estatística E_n .

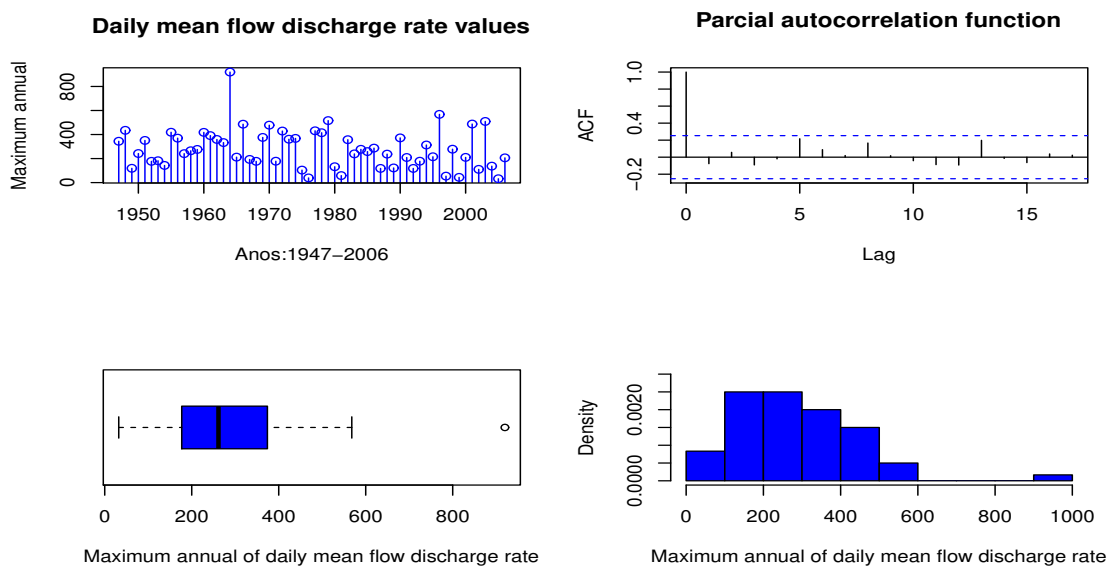


Figura 5.3: Histograma (direita em baixo); caixa-com-bigodes (esquerda em baixo); cronograma (esquerda em cima) e ACF (direita em cima)

apresentam-se as estimativas dos parâmetros da distribuição EV obtidas com o método ML e PWM. São fornecidos também os intervalos de confiança de Log-verosimilhança de perfil.

De salientar que os intervalos de confiança para ξ incluem o zero, o que leva à não rejeição da hipótese nula, $\xi = 0$.

Tabela 5.2: Estimativas dos parâmetros (erros padrões em parênteses) e intervalos de 95% confiança de Log-verosimilhança de perfil (Log-L)

	$\hat{\lambda}$	$\hat{\delta}$	$\hat{\xi}$
ML	210.08 (18.77)	129.81 (13.62)	-0.03 (0.09)
PWM	213.65	137.37	-0.09
	λ	δ	ξ
pLog-V	(174.15;248.23)	(106.41;160.79)	(-0.16;0.19)

5.1.4 A metodologia POT

A aplicação da metodologia POT baseia-se na modelação da distribuição dos excessos acima de um nível elevado, u , na qual se restringe a nossa atenção às observações que excedem esse nível. O desafio desta análise é exatamente a escolha deste limiar. Uma escolha muito elevada pode levar a um número pequeno de observações na cauda, resultando em estimadores com variância elevada, pelo contrário, um limiar baixo, pode levar à violação do Teorema 2.2.6 de Pickands. Esta escolha pode ser feita a partir do gráfico da vida residual média (MRL, do inglês *Mean Residual Life*), baseado no valor médio da distribuição GP , em (2.18), que é uma função linear de u . Se o modelo GP é válido para os excessos acima de um limiar u_0 então será válido também para todos os $u > u_0$. Assim este gráfico deverá evidenciar um comportamento linear acima da escolha adequada do limiar u . Outra metodologia gráfica é baseada no gráfico "Threshold Choice"(TC), que representa as trajetórias das estimativas do modelo GP em função de um conjunto de limiares. O limiar u será uma escolha adequada se as trajetórias das estimativas parecerem ser constantes acima de u . A Figura 5.4 sugere um limiar à volta de 200.

Foi escolhido um limiar $u = 180$, correspondendo a 254 excedências. A Figura 5.5 mostra essas excedências, a não correlação dos dados e alguma assimetria. Foi ajustado o modelo GP aos excessos e obtiveram-se os resultados que constam na Tabela 5.3. Note-se que estes resultados, uma vez mais, indicam um valor de ξ próximo de zero.

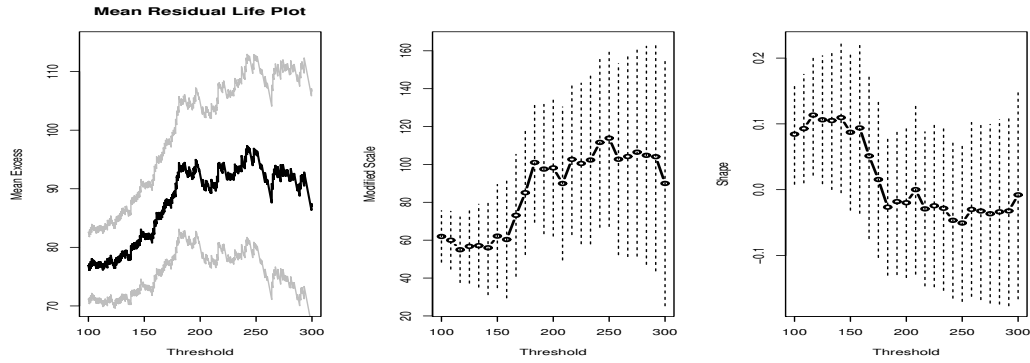


Figura 5.4: Gráfico MRL (esquerda) e gráficos TC (centro e direita)

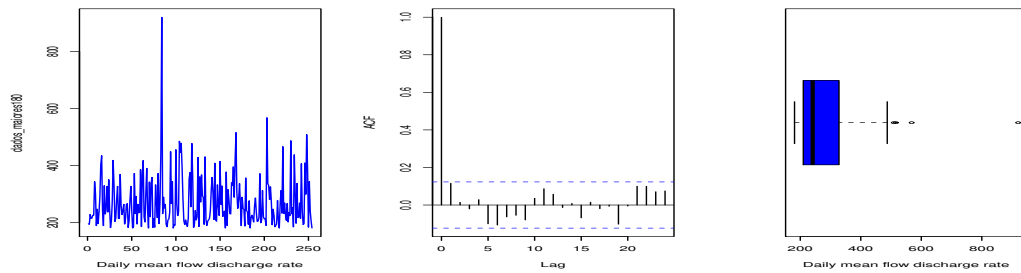


Figura 5.5: Cronograma (esquerda) com $u=180$, função de autocorrelação parcial (centro) e caixa-com-bigodes (direita).

Tabela 5.3: Estimativas dos parâmetros (erros padrões em parênteses) e intervalos de 95% confiança de Log-verossimilhança de perfil (Log-L)

	$\hat{\delta}$	$\hat{\xi}$
ML	94.69 (7.83)	-0.02 (0.05)
PWM	95.47(9.47)	-0.03 (0.07)
	δ	ξ
Log-L	(80.01;110.80)	(-0.10;0.11)

5.1.5 A abordagem semi-paramétrica

Em contexto semi-paramétrico, ξ é o parâmetro primordial em EVT, uma vez que determina a forma da cauda da f.d. F subjacente. As estimativas são baseadas nas k e.o. de topo da amostra, com k uma sucessão intermédia, admitindo que $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})$, para um valor específico de ξ . Uma vez que existem procedimentos de estimação

mais adequados de acordo com o domínio de atração em que se trabalha, deve-se iniciar esta abordagem por testar o max-domínio da Gumbel contra o max-domínio da Fréchet ou da Weibull.

A escolha da cauda

A hipótese a testar é:

$$H_0 : F \in \mathcal{D}_{\mathcal{M}}(EV_0) \quad vs \quad H_1 : F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})_{\xi \neq 0}, \quad (5.5)$$

ou versus as alternativas unilaterais $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})_{\xi < 0}$ ou $F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})_{\xi > 0}$.

Neves and Fraga Alves (2008) apresentaram uma revisão geral dos testes que permitem a escolha do domínio de atração. Em particular, é feito referência aos testes propostos por Neves *et al.* (2006) e Neves and Fraga Alves (2007), em contexto semi-paramétrico, baseados nos k excessos acima da $(n - k)$ -ésima e.o. intermédia ascendente, $X_{n-k:n}$. Em Neves *et al.* (2006) foi introduzida a estatística de teste definida em (5.6). Neves and Fraga Alves (2007) propuseram duas estatísticas de teste: a estatística W_n , em (5.8), baseada numa reformulação da estatística de Hasofer and Wang (1992), e a estatística G_n , em (5.7), do tipo da estatística de teste de Greenwood (Greenwood, 1946).

Assim, sob a hipótese nula do domínio de atração da Gumbel, e admitindo ainda: (i) condições de segunda ordem relativas à cauda direita da f.d. F e (ii) $k \equiv k(n)$ uma sucessão intermédia, tal que $k^{1/2}A(n/k) \rightarrow 0$, quando $n \rightarrow \infty$, onde A é uma função relacionada com a condição de segunda ordem, podem-se definir os seguintes testes:

O teste Ratio

$$R_n^*(k) := \frac{X_{n:n} - X_{n-k:n}}{\frac{1}{k} \sum_{i=1}^k (X_{n-i+1:n} - X_{n-k:n})} - \log k \xrightarrow[n \rightarrow \infty]{d} \Lambda, \quad (5.6)$$

O teste GT

$$G_n(k) := \frac{\frac{1}{k} \sum_{i=1}^k (X_{n-i+1:n} - X_{n-k:n})^2}{\left(\frac{1}{k} \sum_{i=1}^k X_{n-i+1:n} - X_{n-k:n} \right)^2}, \quad (5.7)$$

$$G_n^*(k) = \sqrt{k/4} (G_n(k) - 2) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1),$$

O teste HW

$$W_n(k) := \frac{1}{k} \left[1 - \frac{G_n(k) - 2}{1 + (G_n(k) - 2)} \right], \quad (5.8)$$

$$W_n^*(k) = \sqrt{k/4} (kW_n(k) - 1) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1),$$

onde Λ é uma v.a. com distribuição Gumbel.

A hipótese nula em (5.5) é rejeitada se $T_n^* < \chi_{\alpha/2}$ ou $T_n^* > \chi_{1-\alpha/2}$, onde T^* designa genericamente R^* ou G^* ou W^* e χ_q é o quantil de probabilidade q da correspondente distribuição limite.

Observe-se que, como funções de k , $G_n^*(k)$ e $R_n^*(k)$ tendem a ter uma inclinação concordante com o sinal de ξ e $W_n^*(k)$ discordante.

Se o objetivo for os testes unilaterais, e sendo χ_q o quantil de probabilidade q da correspondente distribuição, as regiões críticas para

Gumbel vs domínio de atração da Weibull são:

$$R_n^*(k) < \chi_\alpha \quad G_n^*(k) < \chi_\alpha \quad W_n^*(k) > \chi_{1-\alpha}. \quad (5.9)$$

Gumbel vs domínio de atração da Fréchet são:

$$R_n^*(k) > \chi_{1-\alpha} \quad G_n^*(k) > \chi_{1-\alpha} \quad W_n^*(k) < \chi_\alpha. \quad (5.10)$$

A Figura 5.6 exhibe as trajetórias amostrais de G^* , R^* e W^* para alguns valores de k . Observa-se que, para a maioria de valores de k , as três estatísticas de teste apresentam valores que pertencem às correspondentes regiões de não rejeição. Assim não existe razão para rejeitar a hipótese nula, $F \in \mathcal{D}_{\mathcal{M}}(EV_0)$.

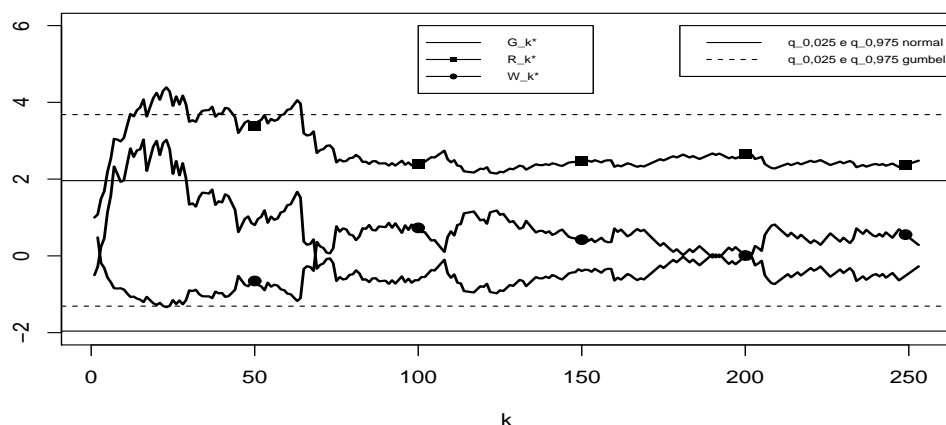


Figura 5.6: Trajetórias amostrais das estatísticas G^* , R^* e W^* .

Algumas estimativas semi-paramétricas

A Figura 5.7 ilustra as trajetórias amostrais das estimativas para cada valor de k . É de salientar que o estimador de Hill e o estimador $CH \equiv MVRB$, especificamente construídos para $\xi > 0$, mostram resultados díspares face aos obtidos com a abordagem paramétrica. Note-se que as estimativas de CH mostram uma trajetória bastante estável, mas próxima de valores positivos de $\hat{\xi}$. Os outros estimadores apresentam trajetórias amostrais próximas de $\hat{\xi} = 0$.

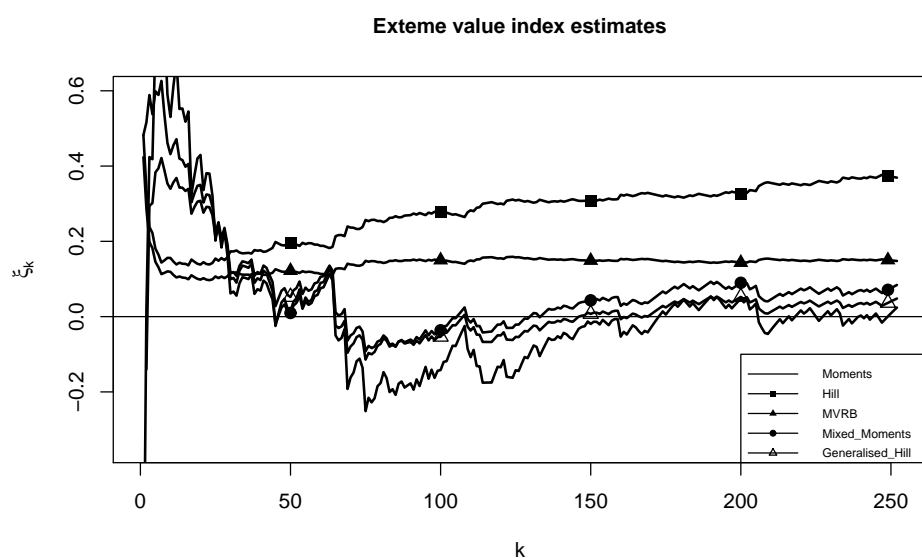


Figura 5.7: Trajetórias amostrais das estimativas de ξ .

5.2 Caso de estudo: área ardida em Portugal

O conjunto de dados utilizado foi o que foi usado no trabalho de Gomes *et al.* (2012) e consiste nos valores diários da área ardida, em Portugal, relativos a cada um dos incêndios existentes nesse dia, durante o período de 1990 a 2003¹, perfazendo um total de 18570 observações.

Para obtenção dos resultados recorreu-se ao *software* R e implementaram-se todos os necessários algoritmos.

No trabalho de Gomes *et al.* (2012) foram considerados como valores extremos todas as observações acima de 100 ha. Assim especificamente os dados consistem nos valores diários de área ardida maiores que 100 ha, relativos a cada um dos incêndios existentes nesse dia, em Portugal, durante o período de 1990 a 2003, perfazendo um total de 2627 observações.

5.2.1 Breve análise preliminar dos dados

Os principais resultados de uma análise gráfica e descritiva preliminar são mostrados na Tabela 5.4 e na Figure 5.8.

Tabela 5.4: Algumas estatísticas descritivas dos dados.

n	Min	1º Quart.	Mediana	Média	3º Quart.	Max	Desv. Padrão	Assim.	Kurt.
2627	100.0340	138.4945	216.412	542.29	430.1155	66070.63	2257.27	20.99	515.86

O histograma, a caixa com bigodes, as estatísticas descritivas, particularmente o coeficiente de assimetria, indicam claramente uma cauda direita pesada.

¹Os dados foram gentilmente cedidos por João Carreiras, a quem se agradece.

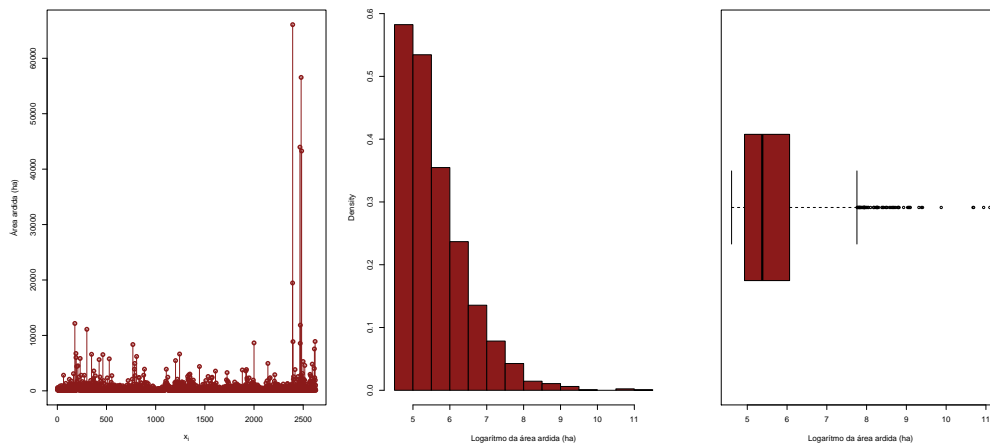


Figura 5.8: Conjunto de dados (esquerda); histograma (centro); caixa-com- bigodes (direita)

5.2.2 Teste à condição de valores extremos

Para o conjunto de dados em análise, testou-se a hipótese em (5.1), através da estatística E_n , em (5.2). A Figura 5.9 mostra a trajetória da estatística E_n e observa-se que todos os seus valores, exceto um, são menores que os correspondentes quantis de probabilidade 0.95. Assim uma vez que a trajetória amostral da estatística de teste se encontra na sua maioria fora da região de rejeição, não existe razão para rejeitar a hipótese nula, em (5.1).

Dietrich, de Haan and Huesler (2002)'s method, eta=2

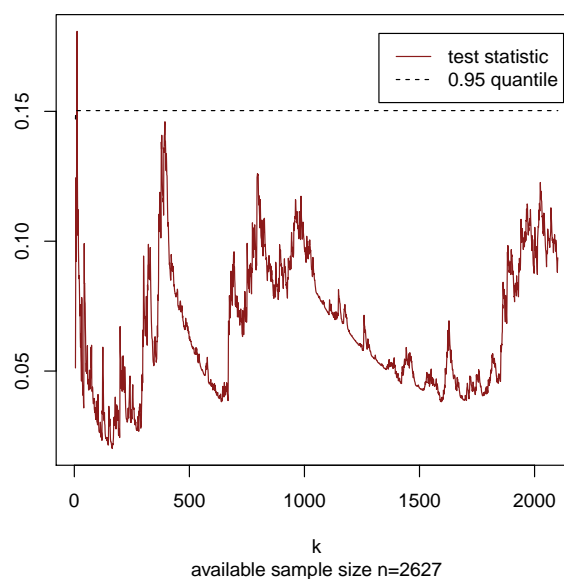


Figura 5.9: Trajetória amostral da estatística E_n .

5.2.3 A escolha da cauda

No âmbito deste trabalho é importante saber se a distribuição do conjunto de dados pertence ao domínio de atração da Fréchet. Assim, a hipótese a testar é:

$$H_0 : F \in \mathcal{D}_{\mathcal{M}}(EV_0)) \quad vs \quad H_1 : F \in \mathcal{D}_{\mathcal{M}}(EV_{\xi})_{\xi>0}. \quad (5.11)$$

A Figura 5.10 exhibe as trajetórias amostrais de G^* , R^* e W^* para alguns valores de k . Observa-se que, as três estatísticas de teste apresentam valores que pertencem às correspondentes regiões de rejeição podendo-se assim concluir pelo domínio de atração da Fréchet.

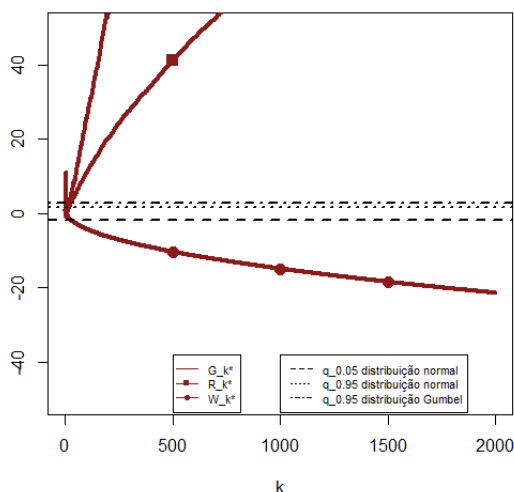


Figura 5.10: Trajetórias amostrais das estatísticas G^* , R^* e W^* .

5.2.4 As estimativas semi-paramétricas em estudo

Considerando que se está a trabalhar no domínio de atração da Fréchet, as classes de estimadores de Lehmer, a classe L_p e a classe L_p^{RB} , assim como a classe GM_p , fazem todo o sentido, como estimadores de EVI. Estes estimadores dependem do valor de p , o qual depende do valor de ρ , como se concluiu na Secção 4.1. O valor que se obteve de $\hat{\rho} \simeq -0.39$ sugere para a classe de estimadores de L_p um valor de $p \simeq 6$ (ver Tabela 4.1) e para a classe de estimadores GM_p , $p \simeq 3$ (ver Tabelas 4.2 e 4.3). A Figura 5.11 mostra as trajetórias amostrais das estimativas para cada valor de k . É de

referir a estabilidade dos estimadores $L_{6.5}$, CH , $L_{1.5}^{RB}$ e $GM_{3.2}$, em particular da classe de estimadores L^{RB} , sugerindo um valor de $\hat{\xi} \simeq 0.65$.

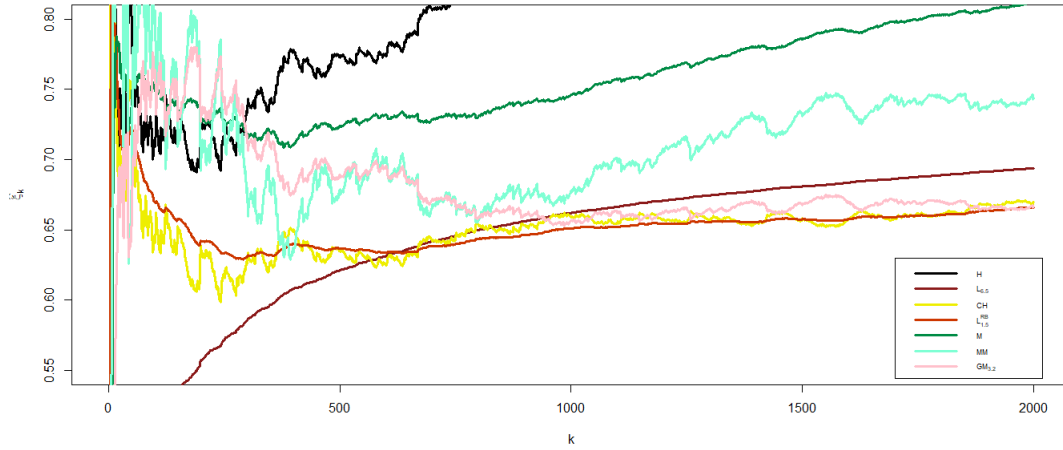


Figura 5.11: Trajetórias amostrais das estimativas de ξ .

Os resultados obtidos são concordantes com os resultados apresentados em Gomes *et al.* (2012) para o mesmo conjunto de dados. A grande estabilidade das trajetórias do estimador L^{RB} é um grande incentivo à pesquisa e construção de procedimentos computacionais para a escolha de p .

Capítulo 6

Conclusões

Neste Capítulo apresentam-se os principais resultados obtidos ao longo deste trabalho de investigação e também os problemas que se deixaram sem resposta e que poderão vir a ser objeto de estudo futuro.

6.1 Resultados

Nesta dissertação foi abordada a estimação, essencialmente em contexto semi-paramétrico, do parâmetro primordial na Teoria de Valores Extremos, o índice de valores extremos, ξ , para modelos de caudas pesadas, i.e, modelos com índice de valores extremos positivo, $\xi > 0$, apresentando alguns contributos computacionais e metodológicos.

Iniciou-se com o enquadramento do tema e com uma breve revisão dos resultados mais relevantes na Teoria de Valores Extremos, considerando que se trabalha com variáveis aleatórias independentes e identicamente distribuídas.

Realizou-se um estudo sobre uma classe de estimadores baseada na média de Lehmer de ordem p de números positivos, que generaliza a média aritmética ($p = 1$) e a média harmónica ($p = 0$). Essa classe de estimadores de Lehmer de ordem p , L_p , constitui uma generalização do clássico estimador de Hill. Dada uma amostra aleatória (X_1, \dots, X_n) e a correspondente amostra das estatísticas ordinais ascendentes

$(X_{1:n}, \dots, X_{n:n})$, o estimador de Hill, pode ser considerado como a média de Lehmer de ordem 1 dos k log-excessos $V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}$, $1 \leq i \leq k < n$. Também foi estudada outra classe de estimadores, a classe de estimadores de Lehmer de ordem p de viés reduzido de segunda ordem, L_p^{RB} , que generaliza um outro estimador de viés reduzido, o estimador de Hill de viés reduzido. Esse estudo consistiu em conhecer o comportamento assintótico, em distribuição, dessas duas classes de estimadores, como também os seus comportamentos em amostras de dimensão finita, procurando efetuar algumas comparações entre elas e com outros estimadores-EVI, que têm constituído investigação recente.

Sob a validade da condição de primeira ordem, a classe de estimadores de Lehmer de ordem p é consistente para a estimação de $\xi > 0$ desde que $p \in \mathcal{R}_L = \{p : p > 0\}$. Sob a validade da condição de segunda ordem e supondo $k = k(n)$ uma sucessão intermédia de inteiros tais que $\lambda_A := \lim_{n \rightarrow \infty} \sqrt{k} A(n/k)$ é um valor finito, é possível escrever-se, para todo o $\rho \leq 0$, e considerando $p > 1/2$, a seguinte representação distribucional,

$$L_p(k) \stackrel{d}{=} \xi + \frac{\sigma_{L_p} Z_k^{(p)}}{\sqrt{k}} + b_{L_p} A(n/k) + o_p(A(n/k)) + o_p(1/\sqrt{k}) \quad (6.1)$$

onde $Z_k^{(p)}$ é uma v.a. com distribuição assintótica normal padrão e com $(b_{L_p}, \sigma_{L_p}^2)$ dados por,

$$b_{L_p} = \frac{1}{(1-\rho)^p} \quad \text{e} \quad \sigma_{L_p}^2 = \frac{\xi^2 \Gamma(2p-1)}{\Gamma^2(p)} \quad (p > 1/2) \quad (6.2)$$

Além disso, se se trabalhar com a classe de modelos Hall-Welsh, e se β e ρ forem estimados consistentemente por $\hat{\beta}$ e $\hat{\rho}$, com $\hat{\rho} - \rho = o_p(1/\ln n)$, consegue-se eliminar a componente dominante do viés de L_p^{RB} , i.e., $b_{L_p^{\text{RB}}} = 0$, e manter a sua variância assintótica igual à do correspondente estimador-EVI de viés não reduzido.

O estudo comparativo assintótico foi realizado com base numa medida de eficiência, genericamente denominada por AREFF, que compara a raiz dos erros quadrados médios assintóticos dos estimadores, no nível ótimo, k_0 , i.e., o nível que minimiza os correspondentes erros quadráticos médios assintóticos. Esse indicador foi calculado para cada um dos estimadores-EVI relativamente ao estimador de Hill, para os quais

essa medida é válida, ou seja apenas contempla os estimadores-EVI que não são de viés reduzido (ou em regiões para as quais o viés assintótico seja diferente de zero). Assim, quanto maior for o valor da medida AREFF, maior a eficiência do estimador face ao estimador de Hill. Conclui-se que o estimador-EVI de Lehmer, no nível ótimo, ultrapassa em eficiência, em todo o plano- (ξ, ρ) , o estimador-EVI H_p , um comportamento atípico entre os estimadores-EVI clássicos. No entanto, o estimador-EVI L , no nível ótimo, pode ser ultrapassado em eficiência pelo estimador-EVI M , ou GH , numa região próxima a $\xi = -\rho/(1 - \rho)$, onde $b_M = b_{GH} = 0$. O estimador-MM, supera em eficiência, nos níveis ótimos, o estimador-M, numa região em redor de $\xi + \rho = 0$, e até superando o estimador-EVI L no nível ótimo. O estimador-EVI GPPWM é um estimador RB para $\xi + \rho = 0$ e pode ser mais eficiente do que o estimador-EVI MM numa pequena região do plano- (ξ, ρ) . De referir ainda que o estimador-EVI PPWM é mais eficiente para valores de ξ em torno de 0.05. Como seria esperado, nenhum dos estimadores domina as alternativas, mas a classe de estimadores-EVI L_p tem um bom desempenho.

Realizou-se um estudo detalhado do comportamento das classes de estimadores L_p e L_p^{RB} , baseado na simulação multi-amostra Monte Carlo, considerando alguns modelos pertencentes à classe de Hall-Welsh. Foram simulados o valor médio e a raiz do erro quadrático médio de todos os estimadores em estudo que se distinguiram no estudo comparativo assintótico nos níveis ótimos e também aqueles que podem ser de viés reduzido, em todo o plano- (ξ, ρ) . Assim para além das classes de estimadores de Lehmer em estudo, consideraram-se nesta análise, o estimador M , MM , H_p e GM_p . As conclusões foram as seguintes:

1. Verifica-se que os valores de p que parecem melhorar o desempenho dos estimadores L e L^{RB} dependem fortemente do valor de ρ . Constata-se que, esses valores são, para $\rho > -1$, consideravelmente maiores do que aqueles que se verificam para $\rho \leq -1$. Especificamente, para $\rho = -0.1$, estes valores situam-se à volta de trinta e para $\rho = -0.5$, à volta de oito. Já quando o $\rho \leq -1$, os valores de p que tornam os estimadores $L_p(k)$ e $L_p^{RB}(k)$ mais eficientes situam-se entre 1.5 e 2.5.
2. Para $\rho > -1$ verifica-se que, independentemente do índice de valores extremos ξ , as trajetórias das estimativas de L e L^{RB} são muito semelhantes, quase

sobrepondo-se, não parecendo haver grande vantagem nos estimadores que são de viés reduzido face aos estimadores correspondentes que não o são.

3. A classe de estimadores L , para valores de p adequados, apresenta uma estabilidade em torno do verdadeiro valor do parâmetro superior quando $\rho < -1$.
4. Quando $\rho = -1$, e para o modelo Fréchet, independentemente do valor de ξ , as trajetórias das estimativas referentes a L^{RB} , para valores de p adequados, exibem uma estabilidade elevada em torno do verdadeiro valor do parâmetro. O mesmo se verifica, para o modelo EV, quando $\xi = 1, 2, 3$.
5. Para $\rho \leq -1$, as trajetórias das estimativas referentes a L^{RB} são sempre superiores às que se observam nos estimadores L . No entanto, para o modelo Burr, verifica-se que a estabilidade das estimativas de L^{RB} se realiza em torno de um valor que é inferior ao valor do parâmetro.
6. De uma maneira geral, as classes de estimadores que parecem revelar uma melhor performance são, para valores de p ótimos:
 - (a) as classes de estimadores L e GM , quando $\rho > -1$;
 - (b) as classes de estimadores L^{RB} e GM , quando $\rho = -1$;
 - (c) as classes de estimadores L e GM , quando $\rho < -1$, apesar da eficiência não ser muito maior quando comparado com o estimador H .

6.2 Trabalho Futuro

No decorrer deste trabalho em vários pontos nos surgiram questões pertinentes que não tendo sido possível abordar nos deixaram vários tópicos que constituem trabalho futuro. Indicamos alguns:

- A utilização de estimadores alternativos de ρ e de β , por considerarmos que a estimação de ξ para os modelos considerados depende fortemente de ρ .
- A escolha de k e de p , para a qual vários procedimentos com recurso a métodos de reamostragem já têm surgido na literatura, mas ainda sem resposta totalmente satisfatória.

- A construção de um *package* em R que incorpore testes à condição de valores extremos e a variedade de estimadores aqui propostos e comparados.

Bibliografia

- [1] Araújo Santos, P., Fraga Alves, M.I. and Gomes, M.I. (2006). Peaks over random threshold methodology for tail index and quantile estimation. *Revstat*, 4(3):227–247.
- [2] Arnold, B., Balakrishnan, N., and Nagaraja, H.N. (1992; 2008). *A First Course in Order Statistics*. 1st Ed., Wiley; 2nd Ed., SIAM.
- [3] Balkema, A.A. and de Haan, L. (1974). Residual life time at great age. *Annals of Probability*, 2:792–804.
- [4] Beirlant, J., Vynckier, P., and Teugels, J.L. (1996). Excess functions and estimation of the extreme-value index. *Bernoulli*, 2(4):293–318.
- [5] Beirlant, J., Dierckx, G., Goegebeur, Y., and Matthys, G. (1999). Tail index estimation and an exponential regression model. *Extremes*, 2(2):177–200.
- [6] Beirlant, J., Goegebeur, Y., Teugels, J., and Segers, J. (2004). *Statistics of Extremes: Theory and Applications*. Wiley, England.
- [7] Beirlant, J., Dierckx, G., and Guillou, A. (2005). Estimation of the extreme-value index and generalized quantile plots. *Bernoulli*, 11(6):949–970.
- [8] Beirlant, J., Caeiro, F., and Gomes, M.I. (2012). An overview and open research topics in statistics of univariate extremes. *Revstat*, 10(1):1–31.
- [9] Beran, J., Schell, D., and Stehlík, M. (2014). The harmonic moment tail index estimator: asymptotic distribution and robustness. *Annals of the Institute of Statistical Mathematics*, 66(1):193–220.
- [10] Bingham, N.H., Goldie, C.M., and Teugels, J.L. (1989). *Regular Variation*. Volume 27, Cambridge University press.

- [11] Brilhante, M.F., Gomes, M.I., and Pestana, D. (2013). A simple generalisation of the Hill estimator. *Computational Statistics & Data Analysis*, 57(1):518–535.
- [12] Brilhante, M.F., Gomes, M.I., and Pestana, D. (2014). The MOP EVI-estimator revisited. In Pacheco, A., Santos, R., Oliveira, M.R., and Paulino, C.D. (eds.), *New Advances in Statistical Modeling and Applications*, pages 163–175. Springer.
- [13] Caeiro, F. (2006). *Estimação de Parâmetros de Acontecimentos Raros*. PhD Thesis, Faculdade de Ciências da Universidade de Lisboa.
- [14] Caeiro, F. and Gomes, M.I. (2002a). Bias reduction in the estimation of parameters of rare events. *Theory of Stochastic Processes*, 8(24):67–76.
- [15] Caeiro, F. and Gomes, M.I. (2002b). A class of asymptotically unbiased semi-parametric estimators of the tail index. *Test*, 11(2):345–364.
- [16] Caeiro, F. and Gomes, M.I. (2006). A new class of estimators of a “scale” second order parameter. *Extremes*, 9(3-4):193–211.
- [17] Caeiro, F. and Gomes, M.I. (2011). Semi-parametric tail inference through probability-weighted moments. *Journal of Statistical Planning and Inference*, 141(2):937–950.
- [18] Caeiro, F. and Gomes, M.I. (2014a). Comparison of asymptotically unbiased extreme value index estimators: A Monte Carlo simulation study. In *AIP Conference Proceedings*, volume 1618, pages 551–554.
- [19] Caeiro, F. and Gomes, M.I. (2014b). A semi-parametric estimator of a shape second-order parameter. In Pacheco, A., Santos, R., Oliveira, M.R., and Paulino, C.D. (eds.), *New Advances in Statistical Modeling and Applications*, pages 137–144. Springer.
- [20] Caeiro, F. and Gomes, M.I. (2015). Bias reduction in the estimation of a shape second-order parameter of a heavy-tailed model. *Journal of Statistical Computation and Simulation*, 85(17):3405–3419.
- [21] Caeiro, F., Gomes, M.I., and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat*, 3(2):113–136.

- [22] Caeiro, F., Gomes, M.I., and Henriques-Rodrigues, L. (2009). Reduced-bias tail index estimators under a third-order framework. *Communications in Statistics—Theory and Methods*, 38(7):1019–1040.
- [23] Caeiro, F., Gomes, M.I., and Vandewalle, B. (2014). Semi-parametric probability-weighted moments estimation revisited. *Methodology and Computing in Applied Probability*, 16(1):1–29.
- [24] Caeiro, F., Gomes, M.I., Beirlant, J., and de Wet, T. (2016a). Mean-of-order p reduced-bias extreme value index estimation under a third-order framework. *Extremes*, 19(4):561–589.
- [25] Caeiro, F., Gomes, M.I., and Henriques-Rodrigues, L. (2016b). A location-invariant probability weighted moment estimation of the extreme value index. *International Journal of Computer Mathematics*, 93(4):676–695.
- [26] Castillo, E., Hadi, A.S., Balakrishnan, N., and Sarabia, J.M. (2005). *Extreme Value and Related Models in Engineering and Science Applications*. Wiley, New Jersey.
- [27] Ciuperca, G. and Mercadier, C. (2010). Semi-parametric estimation for heavy tailed distributions. *Extremes*, 13(1):55–87.
- [28] Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London.
- [29] David, H.A. and Nagaraja, H.N. (2003). *Order Statistics*. 3rd Ed., Wiley.
- [30] Davison, A.C. (1984). Modelling excesses over high thresholds, with an application. In Tiago de Oliveira, J. (ed.), *Statistical Extremes and Applications*, 461–482. D. Reidel.
- [31] Davison, A.C. and Smith, R.L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B. Methodological*, 52(3):393–442.
- [32] de Haan, L. (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*. Mathematical Centre Tract 32, Amsterdam.

- [33] de Haan, L. (1976). Sample extremes: an elementary introduction. *Statistica Neerlandica*, 30(4):161–172.
- [34] de Haan, L. (1984). Slow variation and characterization of domains of attraction. In Tiago de Oliveira, J. (ed.), *Statistical Extremes and Applications*, 31–38. D. Reidel.
- [35] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer Science+Business Media LLC, New York.
- [36] de Haan, L. and Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica*, 52(1):60–70.
- [37] de Wet, T., Goegebeur, Y., and Munch, M.R. (2012). Asymptotically unbiased estimation of the second order tail parameter. *Statistics & Probability Letters*, 82(3):565–573.
- [38] Dekkers, A.L. and de Haan, L. (1993). Optimal choice of sample fraction in extreme-value estimation. *Journal of Multivariate Analysis*, 47(2):173–195.
- [39] Dekkers, A.L., Einmahl, J.H., and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *Annals of Statistics*, 17:1833–1855.
- [40] Deme, E., Gardes, L., and Girard, S. (2013). On the estimation of the second order parameter for heavy-tailed distributions. *Revstat*, 11(3):277–299.
- [41] Dey, D. and Yan, J. (2015). *Extreme Value Modeling and Risk Analysis: Methods and Applications*. Chapman-Hall/CRC.
- [42] Dietrich, D.; De Haan, L. and Hüsler, J.(2002). Testing extreme value conditions, *Extremes*, 5(1):71–85.
- [43] Drees, H. (1996). Refined Pickands estimators with bias correction. *Communications in Statistics-Theory and Methods*, 25(4):837–851.
- [44] Drees, H., Ferreira, A., and de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *The Annals of Applied Probability*, 14(3):1179–1201.
- [45] Drees, H.; De Haan, L. and Li, D. (2006). Approximation to the tail empirical distribution function with application to testing extreme value conditions. *J. Stat. Plan. Inference*, 136, 3498–3538.

- [46] Dwass, M. (1964). Extremal processes. *Annals of Mathematical Statistics*, 35(4):1718–1725.
- [47] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin, Heidelberg.
- [48] Falk, M. (1989). A note on uniform asymptotic normality of intermediate order statistics. *Annals of the Institute of Statistical Mathematics*, 41(1):19–29.
- [49] Feuerverger, A. and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *The Annals of Statistics*, 27(2):760–781.
- [50] Fisher, R.A. and Tippett, L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24:180–190.
- [51] Fraga Alves, M.I. (2001). A location invariant Hill-type estimator. *Extremes*, 4(3):199–217.
- [52] Fraga Alves, M.I., Gomes, M.I., and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica*, 60(2):193–214.
- [53] Fraga Alves, M.I., Gomes, M.I., de Haan, L., and Neves, C. (2009). Mixed moment estimator and location invariant alternatives. *Extremes*, 12(2):149–185.
- [54] Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Polon. Math.*, 6:93–116.
- [55] Geluk, J. and de Haan, L. (1987). *Regular Variation, Extensions and Tauberian Theorems*. Tech. Report CWI Tract 40, Centre for Mathematics and Computer Science, Amsterdam.
- [56] Gnedenko, B. (1943). Sur la distribution limit du terme maximum d'une série aléatoire. *Ann. Math*, 44:423–453.
- [57] Goegebeur, Y., Beirlant, J., and de Wet, T. (2008). Linking Pareto-tail kernel goodness-of-fit statistics with tail index at optimal threshold and second order estimation. *Revstat*, 6(1):51–69.

- [58] Goegebeur, Y., Beirlant, J., and de Wet, T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *Journal of Statistical Planning and Inference*, 140(9):2632–2652.
- [59] Gomes, M.I. (1981). An i -dimensional limiting distribution function of largest values and its relevance to the statistical theory of extremes. In Taillie, C., Patil, G.P., and Baldessari, B.A. (eds.), *Statistical Distributions in Scientific Work*, 6:389–410.
- [60] Gomes, M.I. (1993). On the estimation parameters of rare events in environmental times series. *Statistics for the Environment*, 226–241.
- [61] Gomes, M.I. and Caeiro, F. (2014). Efficiency of partially reduced-bias mean-of-order- p versus minimum-variance reduced-bias extreme value index estimation. In Colubi, A., Blanco, A. & Gatu, C. (eds.), *Proceedings of COMPSTAT 2014*, pages 289–298. The International Statistical Institute/International Association for Statistical Computing.
- [62] Gomes, M.I. and Guillou, A. (2015). Extreme value theory and statistics of univariate extremes: A review. *International Statistical Review*, 83(2):263–292.
- [63] Gomes, M.I. and Henriques-Rodrigues, L. (2010). Comparison at optimal levels of classical tail index estimators: a challenge for reduced-bias estimation? *Discussiones Mathematicae Probability and Statistics*, 30(1):35–51.
- [64] Gomes, M.I. and Henriques-Rodrigues, L. (2016). Competitive estimation of the extreme value index. *Statistics & Probability Letters*, 117:128–135.
- [65] Gomes, M.I. and Martins, M.J. (2001). Generalizations of the Hill estimator— asymptotic versus finite sample behaviour. *Journal of Statistical Planning and Inference*, 93(1):161–180.
- [66] Gomes, M.I. and Martins, M.J. (2002). “Asymptotically unbiased” estimators of the tail index based on external estimation of the second order parameter. *Extremes*, 5(1):5–31.
- [67] Gomes, M.I. and Martins, M.J. (2004). Bias reduction and explicit semi-parametric estimation of the tail index. *Journal of Statistical Planning and Inference*, 124(2):361–378.

- [68] Gomes, M.I. and Neves, C. (2008). Asymptotic comparison of the mixed moment and classical extreme value index estimators. *Statistics & Probability Letters*, 78(6):643–653.
- [69] Gomes, M.I. and Pestana, D. (2007a). A simple second order reduced-bias extreme value index estimator. *J. Statisc. Computation and Simulation*, 77(6):487–504.
- [70] Gomes, M.I. and Pestana, D. (2007b). A sturdy reduced-bias extreme quantile (VaR) estimator. *Journal of the American Statistical Association*, 102(477):280–292.
- [71] Gomes, M.I., Martins, M.J., and Neves, M.M. (2000). Alternatives to a semi-parametric estimator of parameters of rare events—the jackknife methodology. *Extremes*, 3(3):207–229.
- [72] Gomes, M.I., de Haan, L., and Peng, L. (2002). Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes*, 5(4):387–414.
- [73] Gomes, M.I., Caeiro, F., and Figueiredo, F. (2004). Bias reduction of a tail index estimator through an external estimation of the second-order parameter. *Statistics*, 38(6):497–510.
- [74] Gomes, M.I., Pereira, H., and Miranda, M.C. (2005). Revisiting the role of the jackknife methodology in the estimation of a positive tail index. *Communications in Statistics-Theory and Methods*, 34(2):319–335.
- [75] Gomes, M.I., Martins, M.J., and Neves, M.M. (2007a). Improving second order reduced bias extreme value index estimation. *Revstat*, 5(2):177–207.
- [76] Gomes, M.I., Miranda, C., and Viseu, C. (2007b). Reduced-bias tail index estimation and the jackknife methodology. *Statistica Neerlandica*, 61(2):243–270.
- [77] Gomes, M.I., Fraga Alves, M.I., and Araújo Santos, P. (2008a). Port Hill and moment estimators for heavy-tailed models. *Communications in Statistics—Simulation and Computation*, 37(7):1281–1306.
- [78] Gomes, M.I., de Haan, L., and Henriques-Rodrigues, L. (2008b). Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses.

Journal of the Royal Statistical Society: Series B (Statistical Methodology), 70(1):31–52.

- [79] Gomes, M.I., Canto e Castro, L., Fraga Alves, M.I., and Pestana, D. (2008c). Statistics of extremes for iid data and breakthroughs in the estimation of the extreme value index: Laurens de Haan leading contributions. *Extremes*, 11(1):3–34.
- [80] Gomes, M.I., Henriques-Rodrigues, L., Pereira, H., and Pestana, D. (2010). Tail index and second-order parameters' semi-parametric estimation based on the log-excesses. *Journal of Statistical Computation and Simulation*, 80(6):653–666.
- [81] Gomes, M.I., Figueiredo, F., and Neves, M.M. (2012). Adaptive estimation of heavy right tails: resampling-based methods in action. *Extremes*, 15:463–489.
- [82] Gomes, M.I., Fraga Alves, M.I., and Neves, C. (2013a). *Análise de Valores Extremos: uma Introdução*. Edições SPE & INE.
- [83] Gomes, M.I., Martins, M.J., and Neves, M.M. (2013b). Generalized jackknife-based estimators for univariate extreme-value modeling. *Communications in Statistics-Theory and Methods*, 42(7):1227–1245.
- [84] Gomes, M.I., Brilhante, M.F., Caeiro, F., and Pestana, D. (2015a). A new partially reduced-bias mean-of-order p class of extreme value index estimators. *Computational Statistics & Data Analysis*, 82:223–237.
- [85] Gomes, M.I., Martins, M.J., Neves, M.M., and Figueiredo, F. (2015b). Resampling methodologies and reliable tail estimation. *South African Statistical Journal*, 49(1):1–20.
- [86] Gomes, M.I., Brilhante, M.F., and Pestana, D. (2016a). New reduced-bias estimators of a positive extreme value index. *Communications in Statistics-Simulation and Computation*, 45(3):833–862.
- [87] Gomes, M.I., Penalva, H., Caeiro, F. and Neves, M. (2016b). Nonreduced versus reduced-bias estimators of the extreme value index-efficiency and robustness. A. Colubi, A. Blanco and C. Gatu (eds.) *COMPSTAT 2016-22nd International Conference on Computational Statistics*, pages 279–290. The International Statistical/Institute International Association for Statistical Computing.

- [88] Greenwood, M. (1946). The statistical study of infectious diseases. *J. Roy. Statist. Soc. Ser. A*, 109:85–109.
- [89] Greenwood, J.A., Landwehr, J.M., Matalas, N.C., and Wallis, J.R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, 15(5):1049–1054.
- [90] Gumbel, E.J. (1958). *Statistics of Extremes*. Columbia University Press, New York.
- [91] Hall, P. and Welsh, A. (1985). Adaptive estimates of parameters of regular variation. *The Annals of Statistics*, 13(1):331–341.
- [92] Hasofer, A.M. and Wang, Z. (1992). A test for extreme value domain of attraction. *J. Amer. Statist. Assoc.*, 87:171–177.
- [93] Henriques-Rodrigues, L. (2009). *Estimação de Viés Reduzido em Estatísticas de Extremos*. PhD thesis, Faculdade de Ciências da Universidade de Lisboa.
- [94] Henriques-Rodrigues, L., Gomes, M.I., Fraga Alves, M.I., and Neves, C. (2014). PORT-estimation of a shape second-order parameter. *Revstat*, 12(3):299–328.
- [95] Henriques-Rodrigues, L., Gomes, M.I., and Manjunath, B. (2015). Estimation of a scale second-order parameter related to the PORT methodology. *Journal of Statistical Theory and Practice*, 9(3):571–599.
- [96] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3(5):1163–1174.
- [97] Hosking, J., Wallis, J.R., and Wood, E.F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3):251–261.
- [98] Hüsler, J. and Li, D. (2006). On testing extreme value conditions. *Extremes*, 9, 69–86.
- [99] Jenkinson, A.F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158–171.

- [100] Ling, C., Peng, Z., and Nadarajah, S. (2012). Location invariant Weiss-Hill estimator. *Extremes*, 15(2):197–230.
- [101] Neves, C. and Fraga Alves, M.I. (2007). Semi-parametric approach to the Hasofer-Wang and Greenwood statistics in extremes. *TEST*, 16:297–313.
- [102] Neves, C. and Fraga Alves, M.I. (2008). Testing extreme value conditions—an overview and recent approaches. *REVSTAT-Statistical Journal*, 6(1):83–100.
- [103] Neves, C., Picek, J. and Fraga Alves, M.I. (2006). The contribution of the maximum to the sum of excesses for testing max-domains of attraction. *J. Statist. Plann. Inference*, 136(4):1281–1301.
- [104] Paulauskas, V. and Vaičiulis, M. (2013). On the improvement of Hill and some others estimators. *Lithuanian Mathematical J.*, 53:336–355.
- [105] Paulauskas, V. and Vaičiulis, M. (2017). A class of new tail index estimators. *Annals of the Institute of Statistical Mathematics*, 69:661–487.
- [106] Penalva, H., Neves, M., and Nunes, S. (2013). Topics in data analysis using R in extreme value theory. *Metodoloski zvezki*, 10(1):17–29.
- [107] Penalva, H., Nunes, S., and Neves, M. (2014a). Estimação do índice de valores extremos em ambiente R - as abordagens paramétrica e semi-paramétrica. In *Estatística: a Ciência da Incerteza, Atas do XXI Congresso Anual da Sociedade Portuguesa de Estatística*, pages 129–140. SPE.
- [108] Penalva, H., Nunes, S., and Neves, M. (2014b). *Statistical Modeling and Inference in Extremes: Applications with R*, chapter Statistical and Biometrical Challenges: Theory and Applications, Biometrie und Medizinische Informatik - Greifswalder Seminarberichte, pages 281–309. Shaker-Verlag, Aachen.
- [109] Penalva, H., Caeiro, F., Gomes, M., and Neves, M. (2016a). An efficient naive generalization of the Hill estimator - discrepancy between asymptotic and finite sample behaviour. *Notas e Comunicações CEAUL 02/2016*.
- [110] Penalva, H., Nunes, S., and Neves, M. M. (2016b). Extreme value analysis—a brief overview with an application to flow discharge rate data in a hydrometric station in the North of Portugal. *REVSTAT-Statistical Journal*, 14(2):193–215.

- [111] Peng, L. (1998). Asymptotically unbiased estimators for the extreme-value index. *Statistics & Probability Letters*, 38(2):107–115.
- [112] Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 3:119–131.
- [113] R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- [114] Reiss, R.-D. and Thomas, M. (2007). *Statistical Analysis of Extreme Values: With Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhäuser Verlag, Basel-Boston-Berlin, 3rd edition.
- [115] Rényi, A. (1953). On the theory of order statistics. *Acta Mathematica Hungarica*, 4(3-4):191–231.
- [116] Smith, R.L. (1984). Threshold methods for sample extremes. In Tiago de Oliveira, J. (ed.), *Statistical Extremes and Applications*, pages 621–638. D. Reidel.
- [117] Smith, R.L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1):67–90.
- [118] Tiago de Oliveira, J. (1984). *Statistical Extremes and Applications*. D. Reidel, Dordrecht.
- [119] von Mises, R. (1936). La distribution de la plus grande de n valeurs. *Rev. Math. Union Interbalcanique*, 1:141–160. Reprinted in Selected Papers of Richard von Mises, *Amer. Math. Soc.*(1964), 2:271–294.
- [120] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815.
- [121] Worms, J. and Worms, R. (2012). Estimation of second order parameters using probability weighted moments. *ESAIM: Probability and Statistics*, 16:97–113.
- [122] Zhou, C. (2009). Existence and consistency of the maximum likelihood estimator for the extreme value index. *Journal of Multivariate Analysis*, 100(4):794–815.

- [123] Zhou, C. (2010). The extent of the maximum likelihood estimator for the extreme value index. *Journal of Multivariate Analysis*, 101(4):971–983.

Apêndice A

Trajetórias de E e RMSE

Modelo Fréchet(ξ)($\rho = -1$)

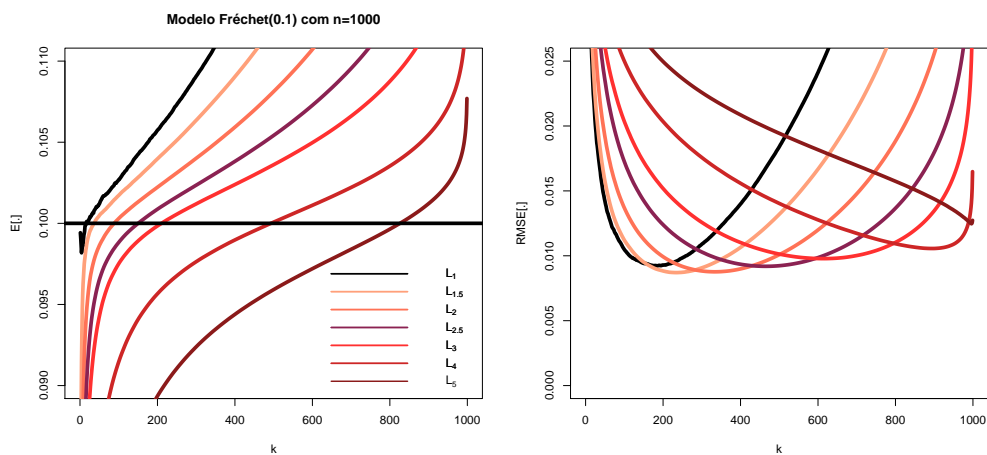


Figura A.1: Trajetórias das estimativas simuladas de E e de RMSE referentes a L .

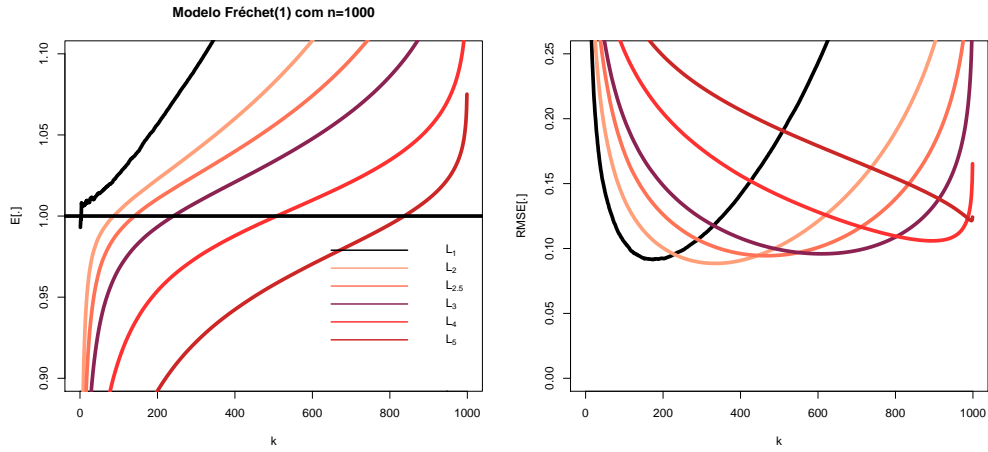


Figura A.2: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1(0.5)3, 4, 5$.

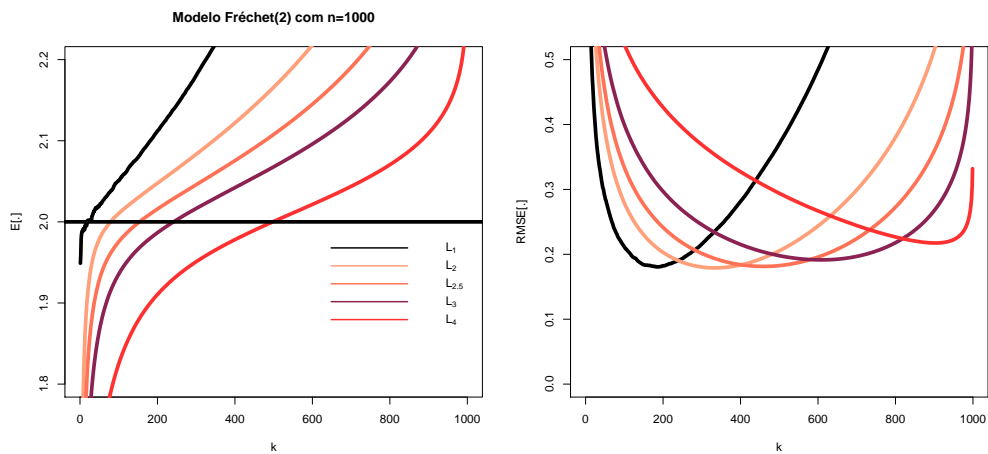


Figura A.3: Trajetórias das estimativas simuladas de E e de RMSE referentes a L .

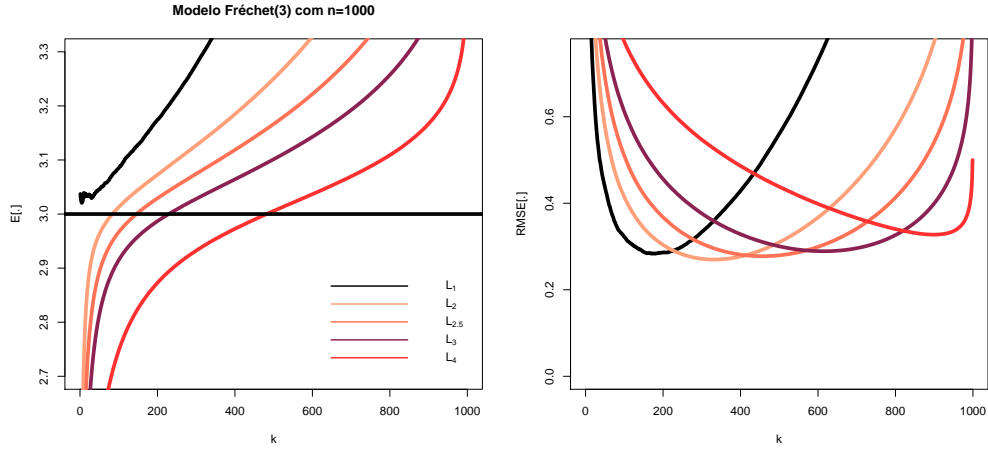


Figura A.4: Trajetórias das estimativas simuladas de E e de RMSE referentes a L.

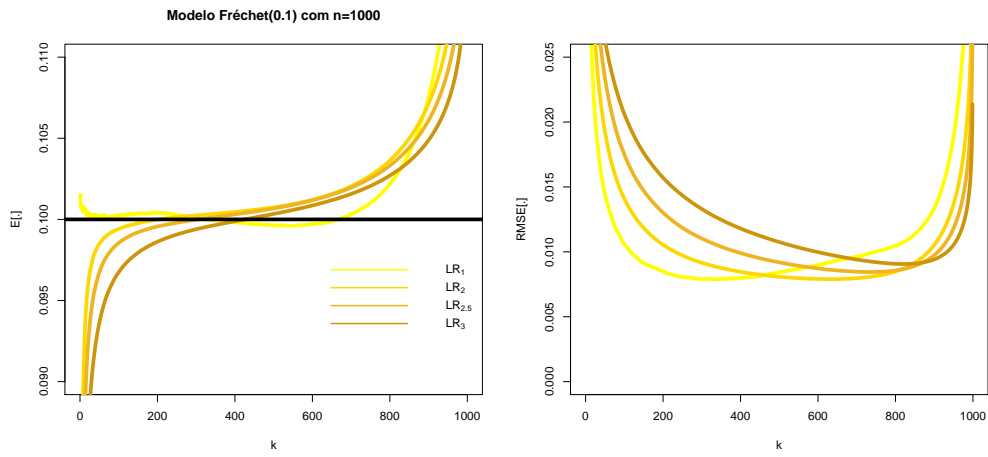


Figura A.5: Trajetórias das estimativas simuladas de E e de RMSE referentes a LR.

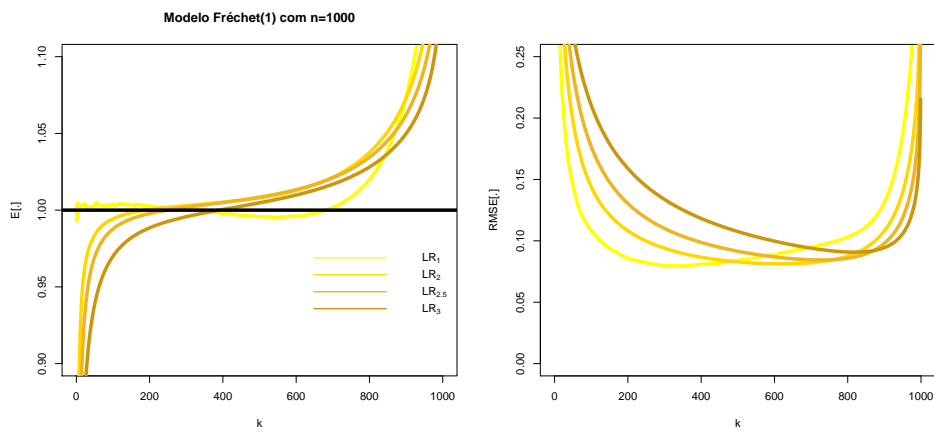


Figura A.6: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{RB} $p = 1, 2, 2.5, 3$.

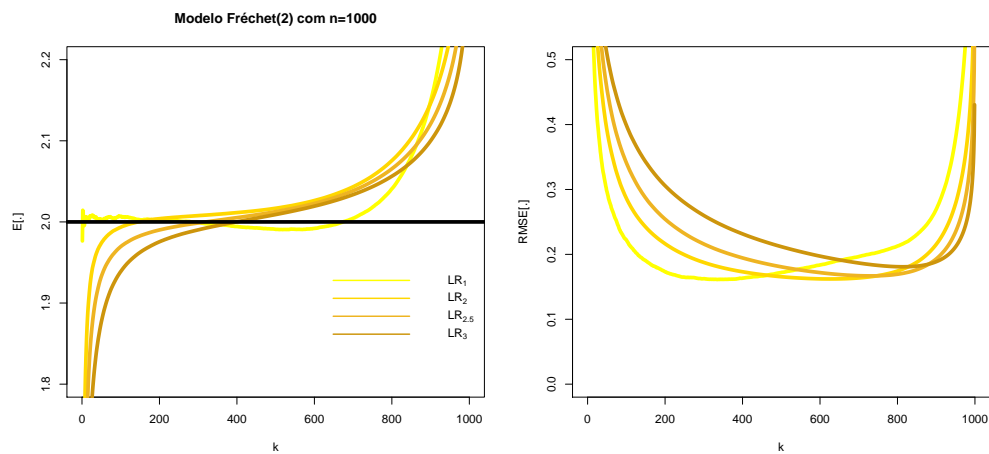


Figura A.7: Trajetórias das estimativas simuladas de E e de RMSE referentes a LR.

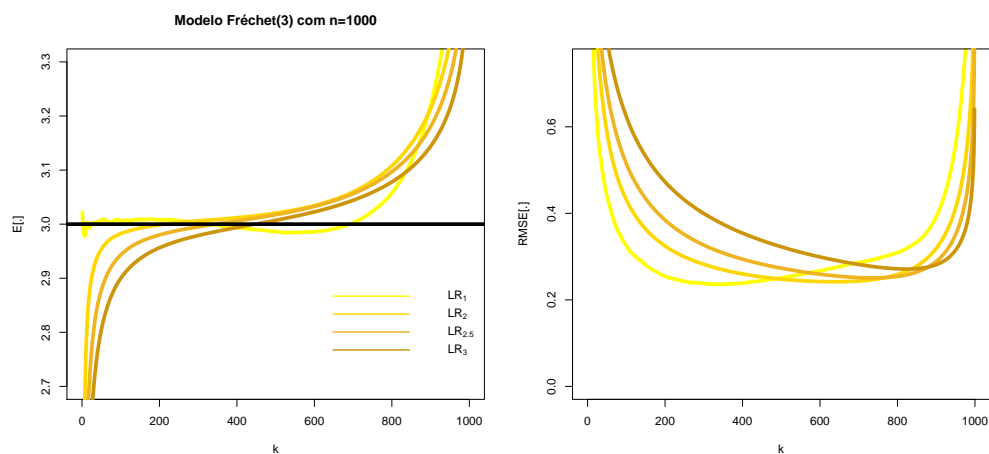


Figura A.8: Trajetórias das estimativas simuladas de E e de RMSE referentes a LR.

Modelo $EV(\xi)(\xi = 1, 2, 3 \text{ e } \rho = -1)$

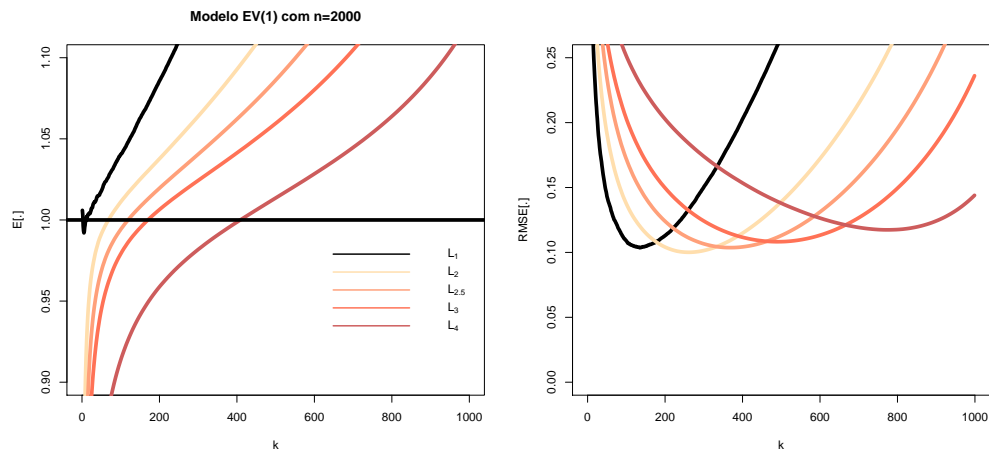


Figura A.9: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 2, 2.5, 3, 4$.

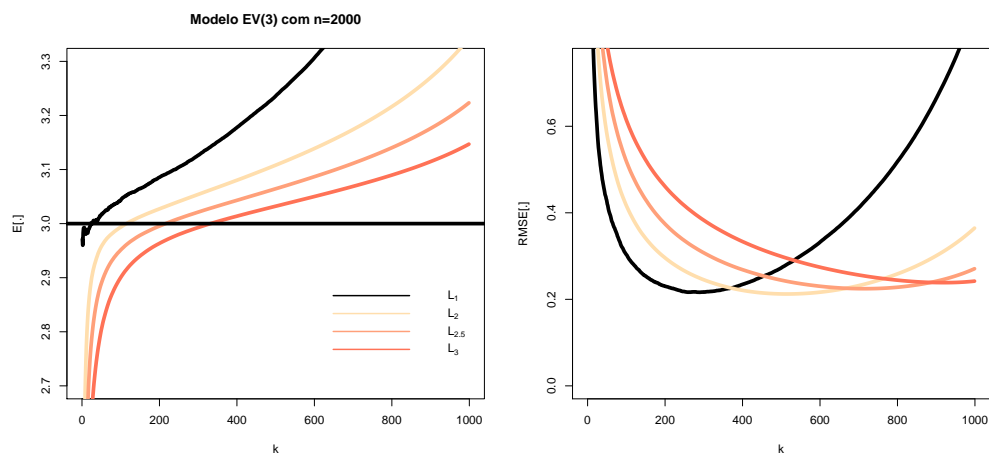


Figura A.10: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 2, 2.5, 3$.

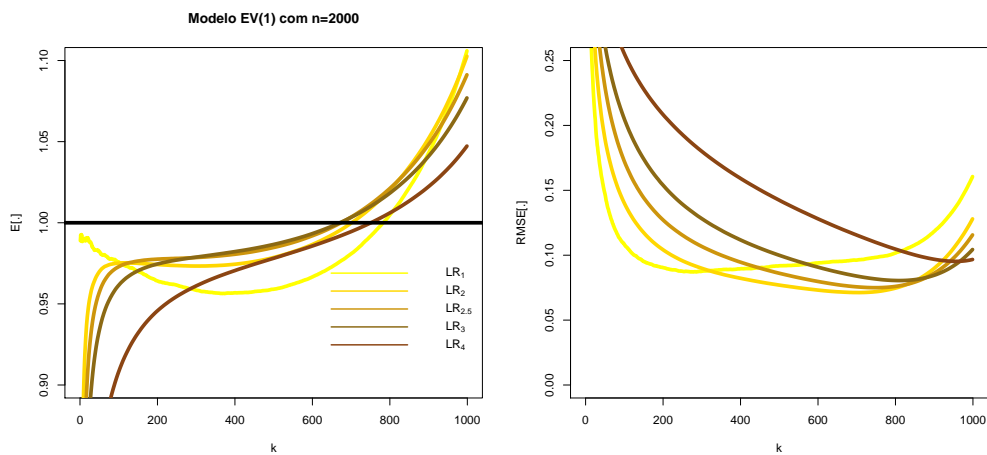


Figura A.11: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{LR} , $p = 1, 2, 2.5, 3, 4$.

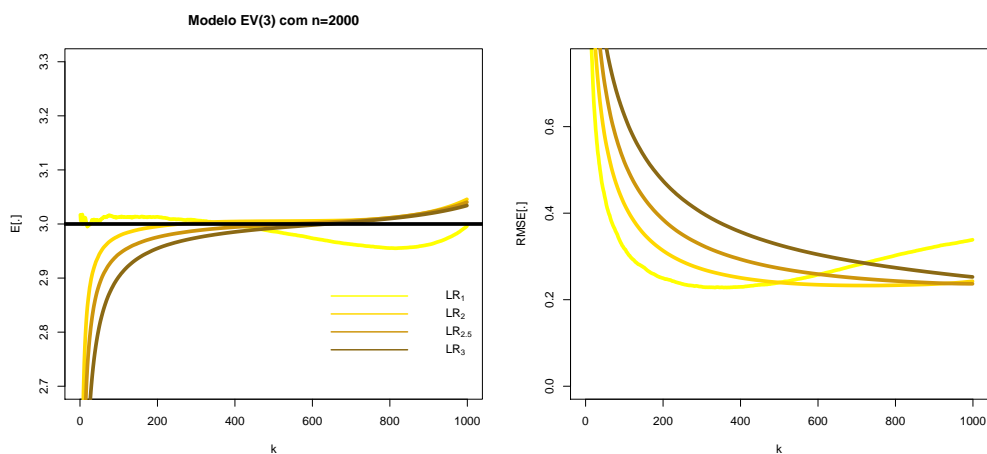


Figura A.12: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{LR} , $p = 1, 2, 2.5, 3$.

Modelo Burr($\xi = 0.5, 1, 2, 3; \rho = -0.1$)

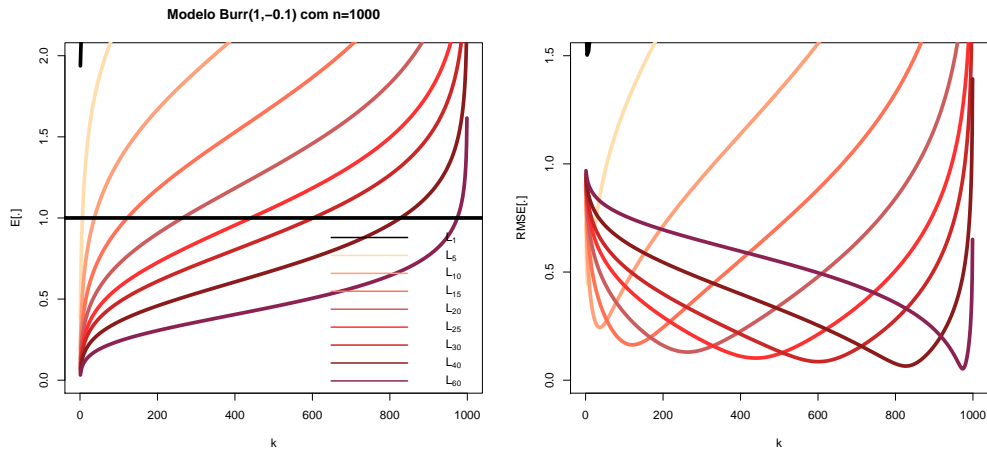


Figura A.13: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 5(5)30, 40, 60$.

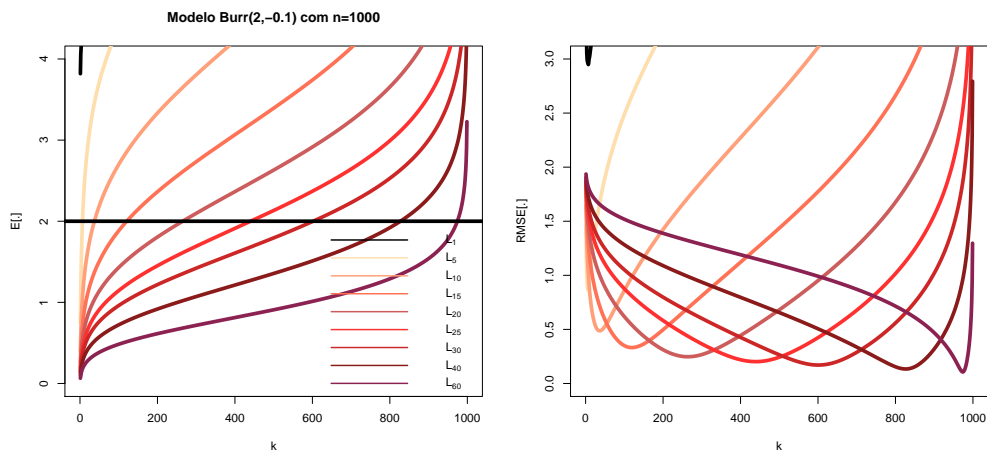


Figura A.14: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 5(5)30, 40, 60$.

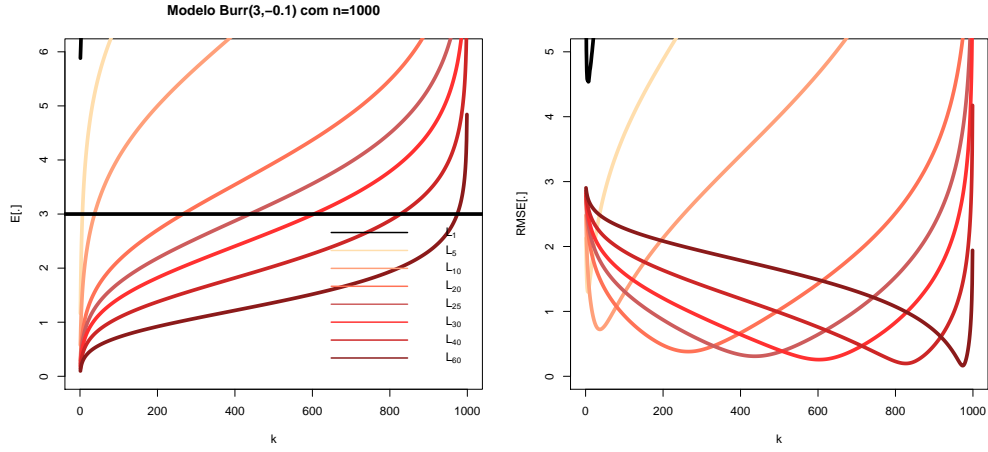


Figura A.15: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 5, 10, 20, 25, 30, 40, 60$.

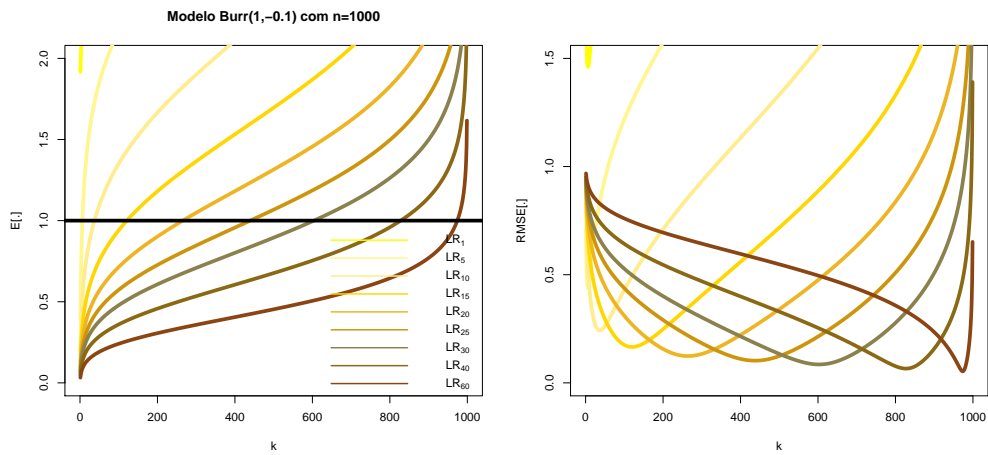


Figura A.16: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{LR} , $p = 1, 5(5)30, 40, 60$.

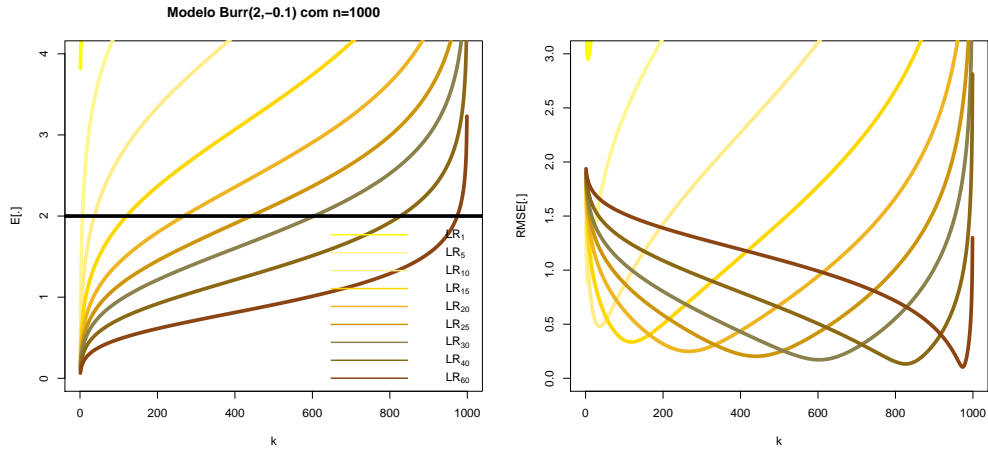


Figura A.17: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{LR} , $p = 1, 5(5)30, 40, 60$.

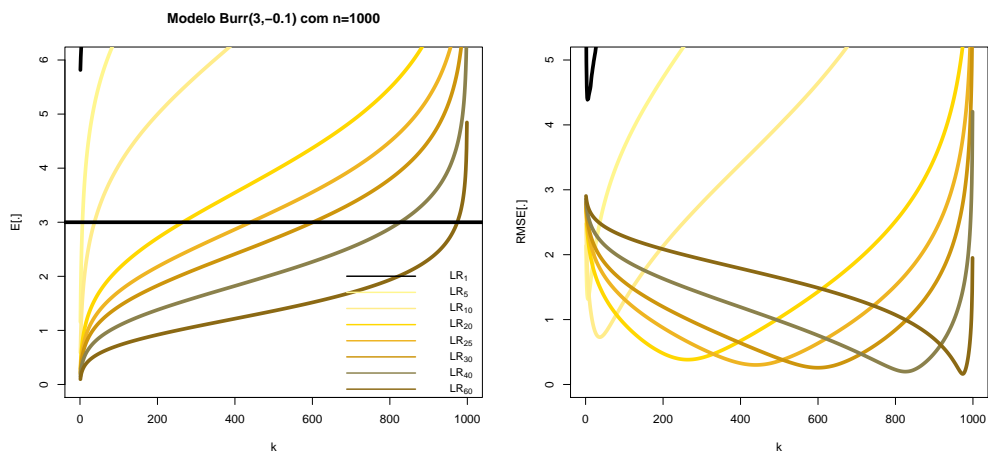


Figura A.18: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{LR} , $p = 1, 5, 10, 20, 25, 30, 40, 60$.

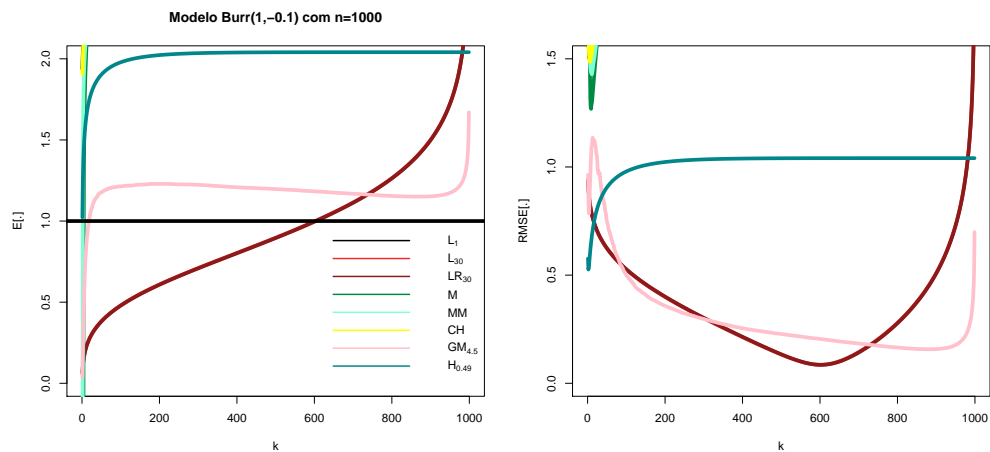


Figura A.19: Comparação das estimativas simuladas de E e de RMSE.

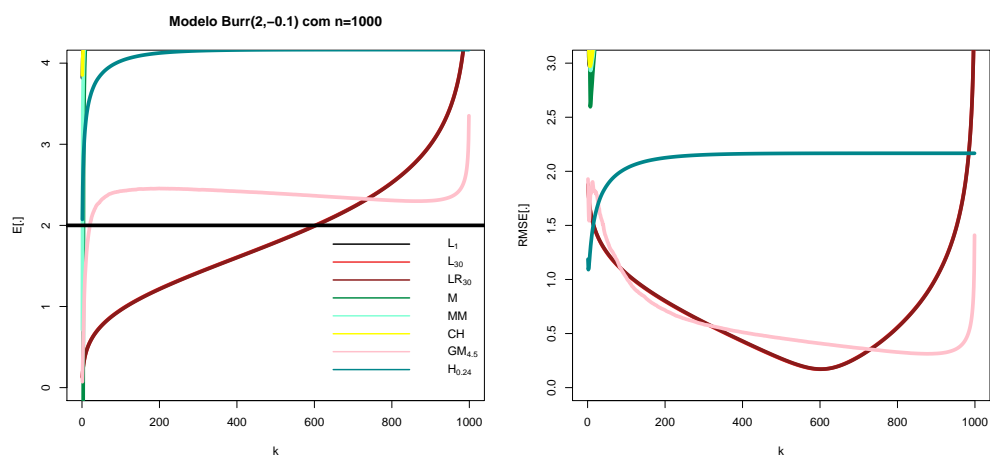


Figura A.20: Comparação das estimativas simuladas de E e de RMSE.

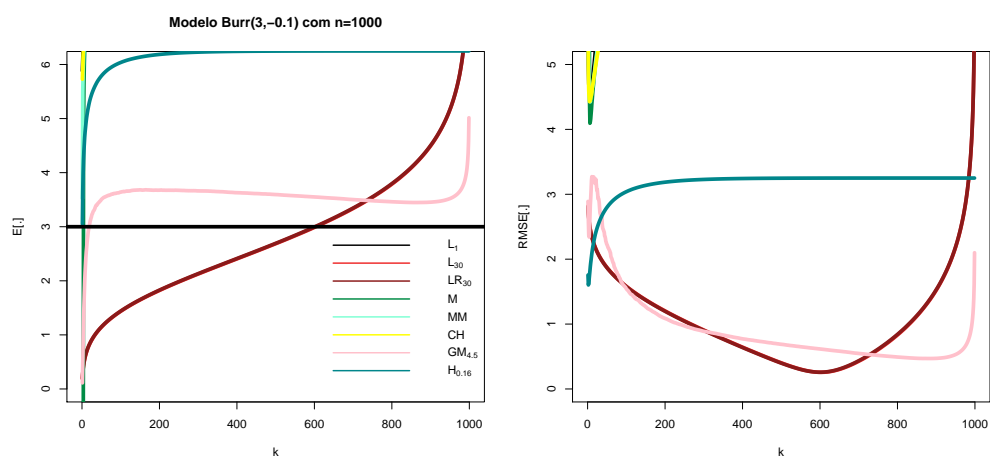


Figura A.21: Comparação das estimativas simuladas de E e de RMSE.

Modelo Burr($\xi = 0.1, 1, 3; \rho = -0.5$)

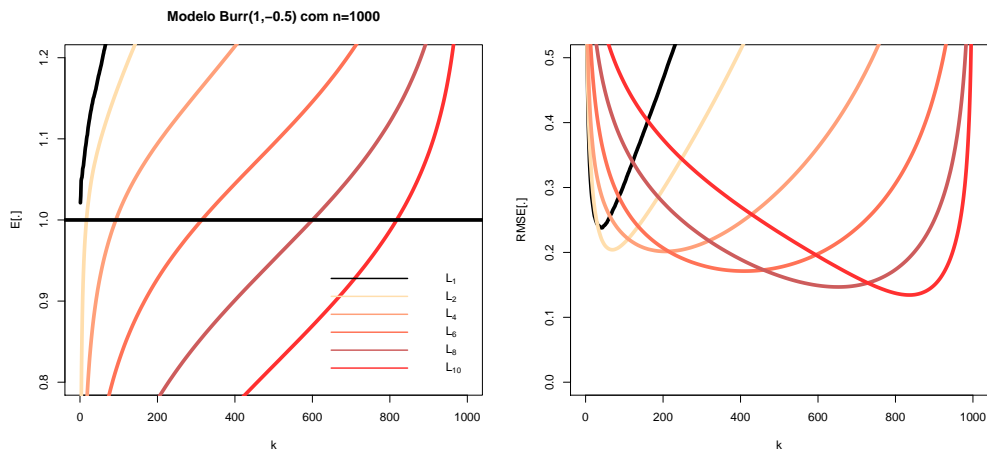


Figura A.22: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 2(2)10$.

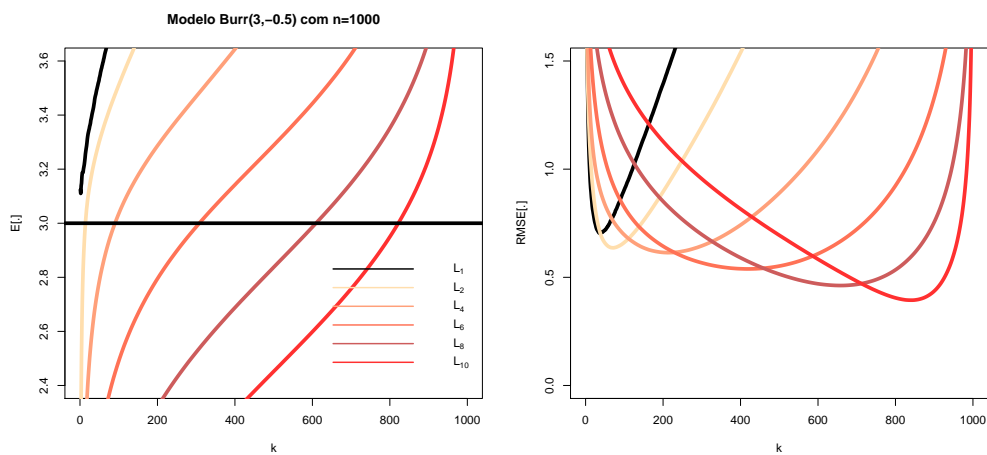


Figura A.23: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1, 2(2)10$.

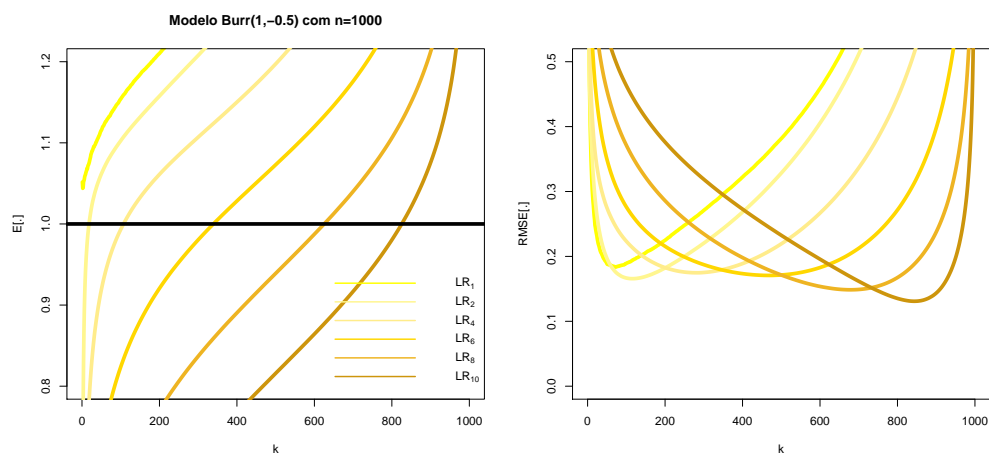


Figura A.24: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{LR} , $p = 1, 2(2)10$.

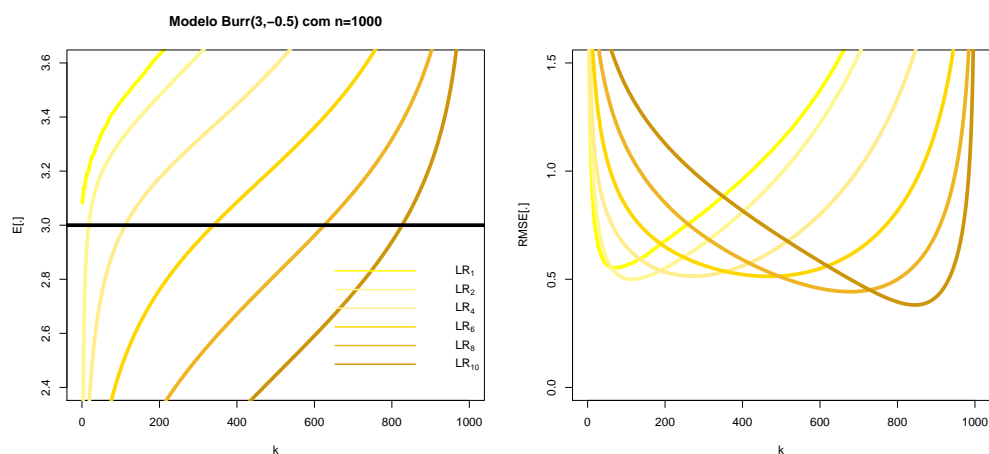


Figura A.25: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p^{LR} , $p = 1, 2(2)10$.

Modelo Burr($\xi = 0.1, 0.5, 1; \rho = -3$)

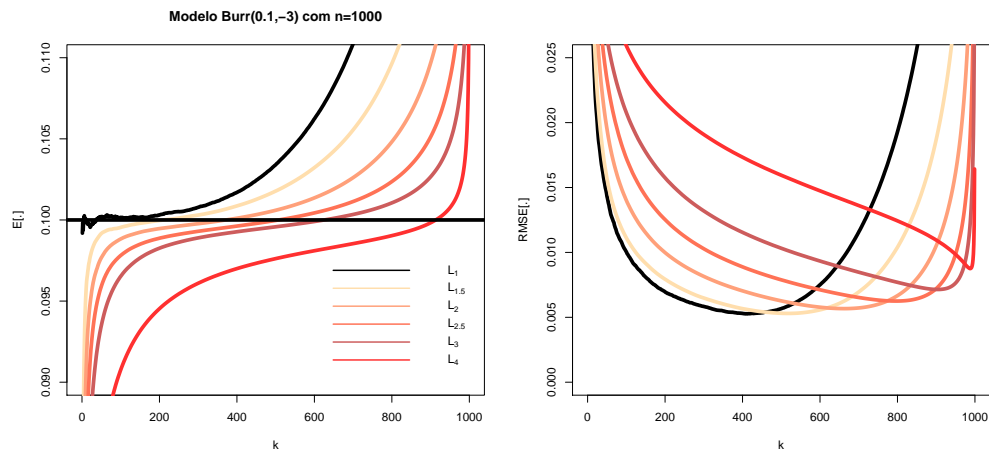


Figura A.26: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1(0.5)3, 4$.

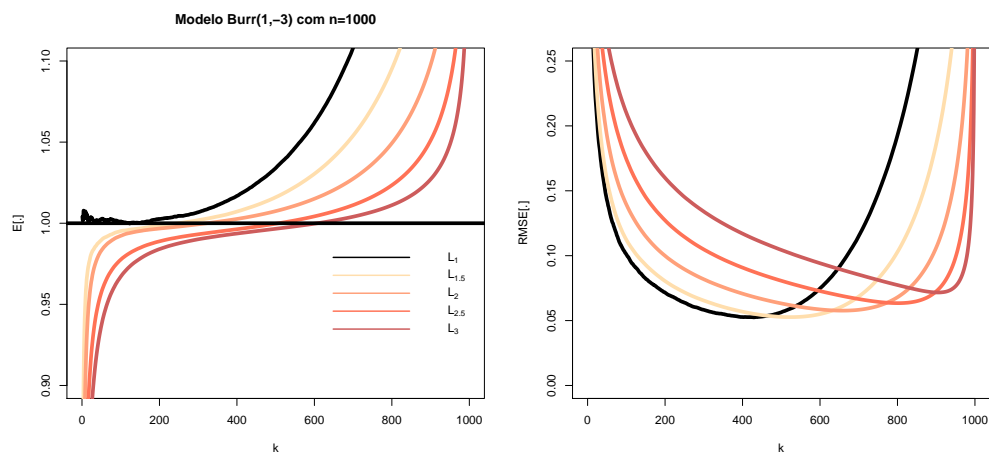


Figura A.27: Trajetórias das estimativas simuladas de E e de RMSE referentes a L_p , $p = 1(0.5)3$.

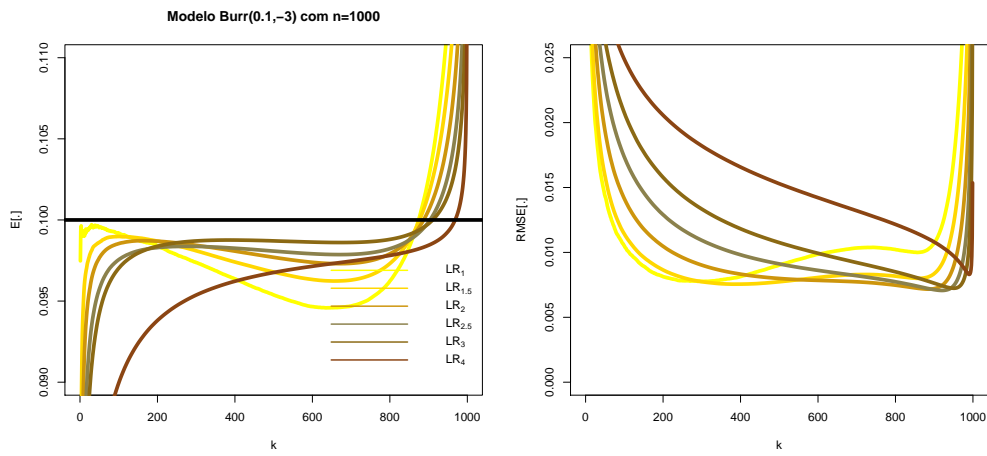


Figura A.28: Trajetórias das estimativas simuladas de E e de $RMSE$ referentes a L_p^{LR} , $p = 1(0.5)3, 4$.

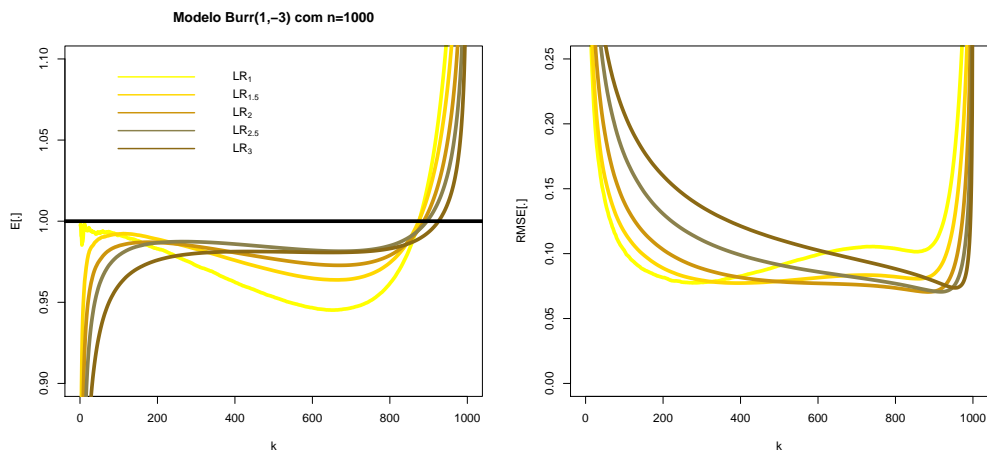


Figura A.29: Trajetórias das estimativas simuladas de E e de $RMSE$ referentes a L_p^{LR} , $p = 1(0.5)3$.